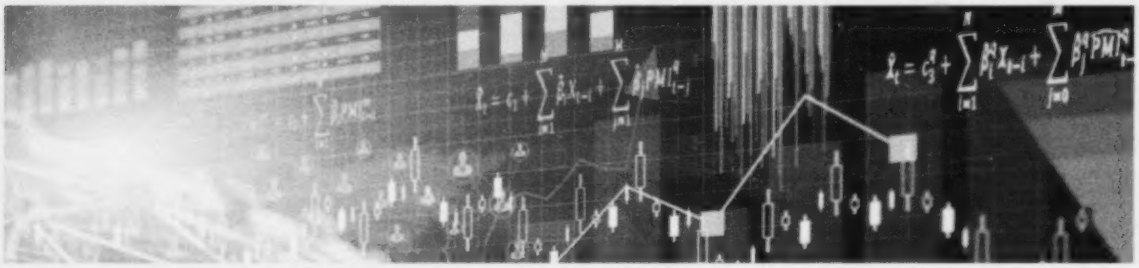
BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel 2016-35

Time-Varying Crash Risk: The Role of Stock Market Liquidity



by Peter Christoffersen, Bruno Feunou, Yoontae Jeon and Chayawat Ornthanalai

Bank of Canada staff working papers provide a forum for staff to publish work-in-progress research independently from the Bank's Governing Council. This research may support or challenge prevailing policy orthodoxy. Therefore, the views expressed in this paper are solely those of the authors and may differ from official Bank of Canada views. No responsibility for them should be attributed to the Bank.

www.bank-banque-canada.ca

Bank of Canada Staff Working Paper 2016-35

July 2016

Time-Varying Crash Risk: The Role of Stock Market Liquidity

by

Peter Christoffersen,[1] Bruno Feunou,[2] Yoontae Jeon[1] and Chayawat Ornthanalai[1]

[1]Rotman School of Management
University of Toronto
Toronto, Ontario, Canada M5S 3E6
peter.christoffersen@rotman.utoronto.ca
yoontae.jeon12@rotman.utoronto.ca
chay.ornthanalai@rotman.utoronto.ca

[2]Financial Markets Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
feun@bankofcanada.ca

Acknowledgements

For helpful comments, we thank Tom McCurdy. This paper has benefited from comments by conference and seminar participants at the Bank of Canada, ESSEC Business School, London Business School (LBS) Doctoral Consortium, Rotman School of Management and University of Geneva. We would like to thank the Bank of Canada, the Global Risk Institute (GRI), and the Social Sciences and Humanities Research Council (SSHRC) for financial support. The authors are responsible for any inadequacies.

Abstract

We estimate a continuous-time model with stochastic volatility and dynamic crash probability for the S&P 500 index and find that market illiquidity dominates other factors in explaining the stock market crash risk. While the crash probability is time-varying, its dynamic depends only weakly on return variance once we include market illiquidity as an economic variable in the model. This finding suggests that the relationship between variance and jump risk found in the literature is largely due to their common exposure to market liquidity risk. Our study highlights the importance of equity market frictions in index return dynamics and explains why prior studies find that crash risk increases with market uncertainty level.

Bank topics: Asset pricing; Financial stability; Econometric and statistical methods
JEL codes: G01, G12

Résumé

Nous estimons un modèle en temps continu à volatilité stochastique et à probabilité dynamique de krach appliqué à l'indice S&P 500. D'après nos conclusions, l'illiquidité du marché supplante les autres facteurs lorsqu'il s'agit d'expliquer le risque d'effondrement boursier. La probabilité de krach est variable dans le temps, mais sa dynamique ne dépend que faiblement de la variance du rendement dès lors que l'illiquidité du marché est intégrée comme variable économique dans le modèle. Au vu de ces résultats, la relation décrite dans la littérature entre la variance et le risque de saut tiendrait en grande partie à leur commune exposition au risque de liquidité du marché. Notre étude souligne ainsi l'importance des frictions du marché des actions dans la dynamique de rendement de l'indice et explique pourquoi les recherches antérieures montrent que le risque de krach augmente avec le niveau d'incertitude du marché.

Sujets de la Banque : Évaluation des actifs; Stabilité financière; Méthodes économétriques et statistiques
Codes JEL : G01, G12

1 Introduction

This paper examines the impact of market liquidity risk on the volatility and crash probability of the aggregate stock market — proxied by the S&P 500 index. We estimate a continuous-time model with stochastic volatility and dynamic crash probability. The innovation of our method is the introduction of market liquidity risk as an economic factor driving the dynamics of volatility and jump intensity. We measure daily market liquidity risk (or "market illiquidity") as the average effective bid-ask spreads of securities constituting the S&P 500 index estimated from high-frequency trades.[1] It can be usefully thought of as the average round-trip trading cost of firms in the index. We estimate the model during 2004–2012 using daily S&P 500 index options, realized spot variance and market illiquidity, and find that 64% of the time-varying crash probability is explained by the stock market's exposure to market liquidity risk.

Market liquidity, defined as the ease with which securities can be bought or sold without significant price impact, has become an increasing concern in financial markets. This is evidenced, for example, by the "flash crash" of May 2010, when major US stock indices fell by almost 10% before recovering quickly. Similarly, market-wide trading halts on August 24, 2015 generated spikes in asset price volatility across financial markets. These two incidents were quickly identified as symptoms of market illiquidity because they occurred in the absence of major news about fundamentals. Unlike the funding liquidity squeeze witnessed in 2007–2008, the current market liquidity risk stems *not* from the banking industry, but perhaps from its absence.[2] In an effort to diminish the chances of repeating the 2008 crisis, regulators and politicians have been working to reduce the role of banks in financial markets, thereby lowering the amount of securities held on bank balance sheets. While this may limit the chances of a subprime crisis repeat, it has the potential to cause investors to increasingly bear the risk from financial markets' trading frictions, e.g., market liquidity risk. As a result, the influence of market liquidity on the economy appears to be increasing in importance.

Market crashes refer to large, unexpected drops in asset prices. Crashes can occur in the presence of information asymmetry about fundamentals, as well in their absence. In the latter case, market liquidity risk is often the culprit. For instance, Huang and Wang (2009) show in an equilibrium framework that when market participation is costly, potential traders are deterred from the market continuously. This causes them to enter the market only when large trading needs arise, which are often on the selling side.[3] Although there exists

[1]This measure is motivated by Aït-Sahalia and Yu (2009) and Goyenko, Holden, and Trzcinka (2009) who find strong empirical supports for using effective bid-ask as the measure for market illiquidity.

[2]Chung and Chuwonganant (2014) find that regulatory changes in the US markets have increased the role of public traders in liquidity provision, which has strengthened the relationship between volatility and market liquidity.

[3]Gennotte and Leland (1990) develop a rational expectation model explaining why a large price drop can occur when there is a relatively small amount of selling in the market. In a more recent study, Cespa and

empirical evidence suggesting that crashes are often driven by market illiquidity, they are typically anecdotal (e.g., "flash crash") or indirect. For instance, Jiang and Yao (2013) find that illiquid stocks have higher daily jump returns in the cross-section. Bradley et al. (2014) find that up to 70% of jumps in equity prices cannot be explained by salient news arrivals, suggesting that trading frictions play an economically important role in causing large stock price changes.

There also exists an extensive literature on index return models that unanimously agrees that index prices "jump."[4] In this case, crashes are large negative jumps in index returns that cannot be explained by the current level of the index's volatility. More recently, several studies have advocated that the probability of observing crashes is time-varying. A typical approach is to let the jump arrival rate increase with the level of the stock return variance.[5] Santa-Clara and Yan (2010) are a notable exception, as they model jump intensity as a quadratic function of state variables. Although this modeling framework is parsimonious, it is inconsistent with the notion that crashes are sudden price drops unexplainable by the current volatility level. Therefore, the relationship between stock market crash risk and market return volatility remains a contentious issue. Further, while recent studies agree that crash risk is time-varying, they are silent on the economic variables driving its dynamic. Our study contributes by providing economic underpinnings to models with time-varying crash risk, and showing that much of the variation in jump intensity is driven by trading frictions, i.e., illiquidity.

To motivate our subsequent modeling framework, we apply a predictive regression analysis linking our market illiquidity measure to a non-parametrically estimated realized jump measure for daily S&P 500 index returns (e.g., Andersen, Bollerslev, and Diebold, 2007; Huang and Tauchen, 2005). Realized jump variation measures the portion of daily return variance that is due to stock price jumps, and we find that it significantly increases with the level of market illiquidity on the previous day. Importantly, the effect of market illiquidity crowds out the predictive ability of realized variance on realized jump variation, suggesting that market illiquidity is the more robust predictor of crash probability for the stock market index. We confirm this finding by running daily time-series regressions on changes in risk-neutral skewness estimated from index option prices, as well as on changes in realized skewness estimated from high-frequency index returns. In either case, we obtain similar conclusions confirming the robust linkage between market illiquidity and crash risk. Armed with this evidence, we estimate a continuous-time model similar to the stochastic volatility with jump model (SVJ) studied by Pan (2002) and Bates (2006), among others. In this model, the jump arrival rate is

Foucault (2014) show that when liquidity providers learn information about an asset from prices of other assets, illiquidity contagion can occur and is a source of market fragility.

[4]This literature is too large to cite in full; for some evidence, see Carr and Wu (2003), Maheu and McCurdy (2004), Andersen, Benzoni, and Lund (2002), Eraker (2004), and Broadie, Chernov, and Johannes (2007).

[5]For examples, see Pan (2002), Bates (2006, 2012), and Christoffersen, Jacobs, and Ornthanalai (2012).

affine in return variance. We extend this framework by letting the time-varying jump intensity dynamic be a function of return variance, market illiquidity, and a latent state variable. We estimate the model by extracting information embedded in index options and high-frequency intraday trades. We use the unscented Kalman filter (UKF) to extract daily state variables. This filtering method allows for sequential learning in the dynamics of latent jump intensity, variance, and illiquidity processes.

We refer to the most general model that we study as the stochastic jump with variance and illiquidity (SJVI). In this model, the jump intensity dynamic, λ_t, is stochastic and affine in the spot variance, the market illiquidity level, and the latent jump-intensity-specific variable designed to captured the omitted risk factor. In all specifications, we model the spot variance as a two-factor square-root process, with market illiquidity being one of the factors. Our estimation results show a strong contemporaneous relationship between market illiquidity and spot variance. On average, a one-standard-deviation increase in the level of market illiquidity increases the spot variance by about 12%. This finding lends support to previous studies that have documented a positive relationship between return volatility and trading activity (e.g., Lamoureux and Lastrapes, 1990; Chae, 2005).

When a jump occurs, its average size is between -3.7% and -5.9% in daily return units. Therefore, jumps represent large drops of daily index price, "crashes," and not market surges. The average jump probability levels are between 2.5 and 2.9 per year with annualized standard deviations between 2.9 and 3.2.

We find strong evidence that during our sample period, crash risk in the S&P 500 index mostly reflects investors' fear of market illiquidity. We arrive at this conclusion by examining the contribution of market illiquidity to the jump intensity dynamic in the new model and find a contemporaneous positive relationship with a strong statistical significance. The influence of market illiquidity dominates other factors, including the market's spot variance. During the 2008 crisis, the influence of spot variance dominates and the contribution of market illiquidity falls to about 30%.

Our results show that market illiquidity is the main economic factor driving crash risk, rather than the level of market's spot variance, which supports our preliminary evidence found using regression analyses. Further, our results suggest the reason previous studies find that jump intensity increases with the level of spot variance is due to the strong positive relationship between variance and market illiquidity.

In terms of economic magnitudes, we find that market illiquidity explains more than half of the S&P 500 index's crash probability level during our sample (64% on average). On the contrary, the contribution of market spot variance to the jump intensity dynamic is only about 12%, with the remaining 24% coming from the latent jump-intensity-specific factor. However, during the six-month period after the Lehman Brothers' collapse in 2008, we find that market

spot variance dominates other factors in explaining the time-varying crash probability, with the contribution as high as 70%. This finding suggests that investors' fear of crash risk during the subprime crisis reflects uncertainty about the market's fundamentals, while outside the crisis period, crash risk mostly reflects investors' fear of market illiquidity.

Estimation of the time-varying jump intensity dynamic is generally difficult, because jumps, and particularly crashes, are rare events (we refer to Bates (2006) for a brief literature review). Therefore, the estimation of time-varying jump dynamics requires econometricians to extract their information from various information-rich sources. Our findings support this view. Using a nested model with fully latent jump intensity dynamic, we find that it produces significantly worse in-sample fits for option prices and spot volatilities. Our study highlights the importance of modeling jump intensity as a function of economic covariates that can be reliably identified from the data, e.g., realized spot variance or market illiquidity.

We emphasize that the relationship between market illiquidity and time-varying volatility and crash risks is not due to market microstructure noise. The market illiquidity proxy that we use is derived from effective spreads of 500 firms constituting the S&P 500 index and *not* from trades on its ETFs nor its futures contracts.

Overall, the findings in this paper emphasize the importance of market liquidity risk in explaining time-varying volatility and crash risks, which is largely missing from prior empirical studies examining index return dynamics. We confirm that our main conclusions hold using various robustness checks. For instance, we show that our estimation results are qualitatively similar before and after implementation of the "circuit breaker" in 2010. We also re-estimate our models using alternative market illiquidity measures including Amihud (2002), and reach the same conclusions.

The remaining parts of this paper proceed as follows. Section 2 describes the data and sample selection, and reports preliminary evidence found using regression analyses. Section 3 describes the model and estimation procedure. Section 4 discusses estimation results and interpretation of our findings. Section 5 demonstrates the robustness of our findings. Finally, Section 6 concludes.

2 Data and Preliminary Evidence

The sample period goes from January 1, 2004 through December 31, 2012. We focus on the recent period because the global financial market has gone through a drastic transformation, e.g., new banking regulations, proliferation of algorithmic trading and exchanged-traded funds. Such changes have strengthened the relationship between market liquidity and stock market volatility, as documented in Chung and Chuwonganant (2014).

2.1 Market Illiquidity

We construct a time-series measure of market liquidity risk at the daily level. We focus on the trading friction associated with the cost of participating in the stock market. We measure this using effective bid-ask spreads following Goyenko, Holden, and Trzcinka (2009), who find strong empirical supports for using intraday bid-ask spreads as the measure of market illiquidity.

We obtain all transactions recorded on securities constituting the S&P 500 index from the TAQ database. Then, for each stock i on day t, we calculate the effective spread of its k^{th} trade as

$$ILQ^i_{t,k} = \frac{2|S^{i,P}_{t,k} - S^{i,M}_{t,k}|}{S^{i,M}_{t,k}}, \tag{1}$$

where $S^{i,P}_{t,k}$ is the price of the k^{th} trade of stock i on day t, and $S^{i,M}_{t,k}$ is the midpoint of the best prevailing bid and ask at the time of the k^{th} trade. The daily effective spread of stock i on day t is then computed as the dollar-volume weighted average effective spreads over all trades during the day[6]

$$ILQ^i_t = \frac{\sum_{k=1}^{K} DolVol^i_{t,k} ILQ^i_{t,k}}{\sum_{k=1}^{K} DolVol^i_{t,k}}. \tag{2}$$

Lastly, we aggregate the effective spreads of firms constituting the S&P 500 index on each day by equally weighting their daily illiquidity measures:

$$ILQ_t = \frac{1}{N} \sum_{i=1}^{N} ILQ^i_t. \tag{3}$$

Daily S&P 500 cash index returns are plotted in the top-left first panel of Figure 1, while the top-right panel plots the daily time series of market illiquidity, *ILQ*. We plot the annualized market illiquidity measure by multiplying their daily levels by 252. The mean annualized *ILQ* measure is 16.85%, which translates to a 0.067% trading cost at the daily level. The standard deviation of the annualized market illiquidity is 5.77% with an interquartile range of 13.90% and 18.26%. We see that the market illiquidity measure rises significantly during the financial crisis period but stays relatively stable in other periods, with occasional few spikes. Figure 1 shows a sharp spike on May 6, 2010, which is associated with the "flash crash" incident.

[6]Figure A.1 in the Appendix shows percentiles of daily dollar effective spread distribution for the S&P 500 constituents. For the majority of firms, their trading cost measured by the dollar effective spread is well above one cent, which is the minimum tick size set by the exchanges. This finding suggests that the effective spread measure that we use is minimally affected by the minimum tick-size rule.

2.2 Realized Variance and Jump Variation

We construct daily realized variance and jump variation measures using intraday S&P 500 cash index returns obtained from TickData. Using the latest observation at each minute, we construct a grid of one-minute intraday returns starting from 9:30 a.m. and ending at 4:00 p.m.

Calculations of realized variance and realized jump variations have been studied extensively in the recent literature.[7] We calculate the daily realized variance measure, *RV*, as the sum of squared one-minute log returns: $RV_t^N = \sum_{i=1}^{N} r_{i,t}^2$. This method measures the total quadratic variation in returns. We measure the variation in daily index returns that is due to the diffusive component using the jump-robust realized variance *MinRV* of Andersen, Dobrev, and Schaumburg (2012). It is calculated as follows:

$$MinRV_t^N = \frac{\pi}{\pi - 2}\left(\frac{N}{N-1}\right)\sum_{i=1}^{N-1}\min(|r_{i,t}|, |r_{i+1,t}|)^2. \tag{4}$$

Following Barndorff-Nielsen and Shephard (2004), we define daily realized jump variation, *RJV*, as the component in total realized variance *RV* that is not explained by *MinRV*. On each day, it is calculated as:

$$RJV_t = \max(RV_t - MinRV_t,\ 0). \tag{5}$$

We can think of *RJV* as the proxy for jump risk in daily index returns. The bottom-left and bottom-right panels in Figure 1 plot the annualized daily time series of *MinRV* and *RJV*, respectively. The mean and standard deviation of the *MinRV* measure over this period are 2.36% and 6.55%, respectively. For the *RJV* measure, the mean and standard deviation are 0.33% and 0.94%. The daily index return variance is thus mostly composed of the continuous component of stock price change.[8]

2.3 Predicting Realized Jumps

This section provides preliminary evidence on the economic relationship between market illiquidity and jump risk. We estimate a predictive regression model on the realized jump variation measure *RJV*. The objective is to identify the economic variables that robustly predict the occurrence and magnitudes of jumps the next day.

We examine three variables of interest and their various combinations, namely, the market

[7] See Huang and Tauchen (2005) for a concise summary.

[8] This finding is consistent with Huang and Tauchen (2005) who find that jumps account for 7% of stock market price variance.

illiquidity measure ILQ, the diffusive quadratic variation measure $MinRV$, and the option-implied volatility index VIX. Most general predictive regression specification is

$$RJV_{t+1} = \beta_0 + \beta_1 MinRV_t + \beta_2 ILQ_t + \beta_3 VIX_t + \beta_3 Ret_t + \beta_3 RJV_t + \epsilon_{t+1}. \quad (6)$$

We obtain daily VIX levels from the Chicago Board of Options Exchange (CBOE) that represent the market's fear index calculated as the 30-day implied volatility level of S&P 500 index options. We include lagged log return of the S&P 500 (Ret) and the autoregressive term for RJV as control variables. Year and day-of-the-week fixed effects are present in all regression specifications, but their estimates are not reported here to save space.

Table 1 summarizes the regression results for six specifications based on the general model described in equation (6). We report the heteroskedasticity-consistent t-statistic (White, 1980) in parentheses below each parameter estimate. Columns (1)–(3) show that when each of the three variables enters in the regression model, it appears statistically significant in predicting the realized jump variation the next day. We find that $MinRV_t$, ILQ_t, and VIX_t are positive and highly significant at the 95% level or higher. These positive coefficients confirm the intuition that jumps are more likely to occur following a day of more volatile and illiquid market conditions. Looking at the size of the coefficients, we find that market illiquidity, ILQ, is the dominant variable. A one-standard-deviation increase in market illiquidity today would increase realized jump variation by 9.1% the next day.

Columns (4)–(5) report results for combining market illiquidity and the two volatility measures. The results clearly show that ILQ_t is the dominant variable in predicting daily realized jump variation. When the market illiquidity measure ILQ_t is added to the regression, the coefficients on $MinRV_t$ and VIX_t variables lose statistical significance. Column (6) reports results for the most general specification where all independent variables are included. We find that the coefficient on ILQ_t decreases by half due to influences of the two market variance measures. Nevertheless, Column 6 shows that ILQ_t is the only variable that remains statistically significant, confirming that it is the leading predictor of realized jump variation in the stock market index.

We emphasize that the market illiquidity measure that we use is calculated from effective spreads of 500 firms constituting the S&P 500 index while the dependent variable, RJV, is constructed from one-minute log returns on the S&P 500 cash index. Market illiquidity and realized return jump variations are thus not related. Further, our finding that market illiquidity dominates return variance in explaining the time-varying market jump risk is confirmed using both $MinRV_t$ and VIX_t.

Overall, the results in Table 1 provide preliminary evidence for the importance of market illiquidity in explaining time-varying jumps in index returns. Further, it shows that omitting

market liquidity risk can lead to a different conclusion regarding the role of market return variance in explaining jumps on the stock market index.

2.4 Market Illiquidity and Crash Risk

The realized jump variation *RJV* measure that we used in the previous section captures the magnitude of positive and negative jumps in index returns and therefore does not identify a stock market crash from a stock market surge. We provide further evidence linking the role of market illiquidity to crash risk by estimating the impact of *ILQ* on daily skewness of the S&P 500 index. Because crashes are large sudden drops in asset prices, a more negative skewness measure would signal a higher probability of crash risk (Chen, Hong, and Stein, 2001). Therefore, if market illiquidity is strongly linked to the stock market crash, we expect that ILQ_t would be negatively related to the stock market skewness measure.

We follow the nonparametric method developed by Bakshi and Madan (2000) and calculate daily skewness from S&P 500 index options. We refer to this measure as risk-neutral skewness, *RNSkew*. We obtained end-of-day S&P 500 index option prices from OptionMetrics. *RNSkew* is calculated from option contracts with approximately one month to maturity. Therefore, the measure $RNSkew_t$ on day t that we use represents investors' forward-looking risk-neutral expectation of the stock market crash risk from the end of day t to day $t + 30$. We discuss details on the skewness measure in the Appendix A.

We examine the impact of market illiquidity risk on the stock market's risk-neutral skewness using the following regression model:

$$\Delta RNSkew_t = \beta_1 \Delta MinRV_t + \beta_2 \Delta ILQ_t + \sum_{i=1}^{p} \alpha_i \Delta RNSkew_{t-i} + \sum_{j=1}^{q} \delta_j \epsilon_{t-j} + \epsilon_t, \quad (7)$$

where Δ indicates that we are examining the change in daily variables, and ϵ_t is normally distributed error term. We estimate the model on change in daily skewness and not on its level because $RNSkew_t$ is highly persistent with an autocorrelation of 0.98. As a result, the explanatory variables that we use are changes in daily *MinRV* and *ILQ*. The regression shown in equation (7) is an autoregressive-moving-average model (ARMA) with p-order lag in the autoregressive term and q-order lag in the error term. Although not shown in equation (7), we include *Ret*, and day-of-the-week fixed effects as control variables in the regression model. We estimate the model using maximum likelihood.

Table 2 reports estimation results for four regression specifications based on the general model in equation (7). We choose the ARMA model with lags of $p = 2$ in the autoregressive term and $q = 1$ in the error term. The number of lags is determined based the LjungBox test corresponding to the most parsimonious model that sufficiently removes autocorrelations in

the residuals. Column (1) provides the baseline regression results. It shows that the change in risk-neutral skewness is negatively related to the index return and is strongly explained by its autoregressive terms. Columns (2) and (3) show that when $\Delta MinRV$ and ΔILQ are separately included in the regression, they load negatively on the change in risk-neutral skewness, although, ΔILQ is the only statistically significant variable. The negative and highly significant coefficient on ΔILQ suggests that when the average trading cost in the stock market increases, investors' expectation of the market crash risk also increases.

Column 4 reports results for the full regression model. We observe two striking findings. First, the coefficient on $\Delta MinRV$ switches sign from negative to positive and is significant at the 10% level. Second, the coefficient on ΔILQ remains negative and significant but approximately doubles in term of magnitude. These results suggest that both $MinRV$ and ILQ are important determinants of daily stock market skewness, i.e., crash risk. The positive coefficient on $\Delta MinRV$, however, shows that an increasing market variance is related to a less negatively skewed risk-neutral distribution of daily index returns. This finding is intuitive because as the variance level increases, the index return distribution would become more fat-tailed on both positive and negative sides, indicating an increasing likelihood of market surges as well as market crashes. Consequently, the index return distribution appears more symmetric, i.e., less negatively skewed, conditional on an increasing variance level. As a robustness check, we verify this result by replacing $MinRV$ with the total quadratic variation (RV) in Table 2 and obtain the same conclusion.

For comprehensiveness, we verify our results using realized skewness constructed from high-frequency trades on the S&P 500 constituents. The method that we use follows that in Amaya, Christoffersen, Jacobs, and Vasquez (2015). We obtain the same conclusion when using daily realized skewness. To save space here, the results are reported in Appendix C, which shows that overall market illiquidity is a leading determinant of the stock market crash risk and that its influence dominates the impact of market return variance. Motivated by this non-parametric evidence, we develop a continuous-time model that allows market illiquidity to act as an economic covariate in explaining the time-varying volatility and crash risks.

3 Model and Estimation

3.1 The SJVI Model

We begin by specifying the processes governing the log stock price, spot variance, spot illiquidity, and latent component of jump intensity dynamic under the risk-neutral measure ($\mathbb{Q}$). We use the notation S_t and V_t to denote stock price and spot variance at time t. We let L_t represent the spot market illiquidity, which measures the liquidity risk of the stock market at

time t, with a higher value indicating a more illiquid market. We include a stochastic process Ψ_t that is designed to capture the latent time-varying jump intensity in index returns. Thus, the model consists of four factors that fully describe the return dynamics under $\mathbb{Q}$:

$$d\log(S_t) = (r - \frac{1}{2}V_t - \xi\lambda_t)dt + \sqrt{V_t}(\sqrt{1-\rho^2}dW_t^1 + \rho dW_t^2) + q_t dN_t \quad (8)$$

$$dV_t = \kappa_V(\theta_V - V_t)dt + \gamma dL_t + \xi_V\sqrt{V_t}dW_t^2 \quad (9)$$

$$dL_t = \kappa_L(\theta_L - L_t)dt + \xi_L\sqrt{L_t}dW_t^3 \quad (10)$$

$$d\Psi_t = \kappa_\Psi(\theta_\Psi - \Psi_t)dt + \xi_\Psi\sqrt{\Psi_t}dW_t^4, \quad (11)$$

where r denotes the risk-free rate and all Brownian motions dW_t^i, for $i = 1$ to 4, are independent of each other.

We assume the market illiquidity process L_t, and the latent jump intensity process Ψ_t, in equations (10) and (11) follow the standard square-root model with long-run mean levels of θ_L and θ_Ψ, respectively. The variance dynamic in equation (9) follows Heston's (1993) square-root process with an additional term γdL_t. We discuss our specification choice for the variance dynamic later in this subsection.

The log stock price dynamic described in equation (8) follows a standard jump-diffusion process where $q_t dN_t$ denotes the jump component. Following the extant literature on index return models, we assume that jumps follow a compound Poisson process with intensity λ_t and each individual jump is independent and identically distributed (i.i.d.) normal with the jump mean size θ and the jump size standard deviation δ. To ensure the discounted log stock price is a martingale, we include the jump compensation term $\xi = e^{(\theta + \frac{\delta^2}{2})} - 1$ in equation (8). Lastly, to complete the model, we specify the dynamic of the time-varying jump intensity λ_t as follows:

$$\text{SJVI model:} \qquad \lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t. \quad (12)$$

This jump intensity specification is motivated by numerical tractability and ease of economic interpretation. Equation (12) shows that the time-varying jump arrival rate is determined jointly by the levels of spot variance V_t, spot market illiquidity L_t, and state variable Ψ_t. The latent state variable Ψ_t is designed to capture the portion of jump intensity dynamic not explained by the covariates V_t and L_t. For the remaining parts of this paper, we refer to this general specification as the SJVI model.

Equation (9) shows that the evolution of spot variance depends on its own mean-reverting drift, the diffusive component, and the market illiquidity process L_t. This specification allows for changes in the market spot variance V_t and market illiquidity L_t to be contemporaneously related, which is supported by Lamoureux and Lastrapes (1990) who find that daily trad-

ing volume significantly explains daily return variance. We choose a parsimonious modeling framework that lets dL_t enter the dynamic of dV_t, and not vice versa. Besides parsimony, this choice is motivated by their joint time-series estimates, which show that the change in market illiquidity leads the change in market spot variance.

This is shown in the Vector Autoregressive Moving-Average (VARMA) model below:

$$\begin{pmatrix} \Delta MinRV_t \\ \Delta ILQ_t \end{pmatrix} = \delta + \Phi \begin{pmatrix} \Delta MinRV_{t-1} \\ \Delta ILQ_{t-1} \end{pmatrix} - \Theta u_{t-1} + u_t, \tag{13}$$

where δ is a 2×1 vector of coefficients, and u_t is a 2×1 vector of normally distributed residuals. Φ and Θ are 2×2 matrices of VARMA model coefficients. The variables $MinRV$ and ILQ are defined in the previous section. We estimate the model above and find that

$$\Phi = \begin{pmatrix} -0.149^{***} & 0.479^{***} \\ (-3.59) & (4.50) \\ -0.023 & 0.237^{***} \\ (-1.47) & (5.45) \end{pmatrix} \quad \text{and} \quad \Theta = \begin{pmatrix} 0.717^{***} & -0.031 \\ (19.97) & (-0.39) \\ -0.014 & 0.730^{***} \\ (-1.18) & (22.09) \end{pmatrix}, \tag{14}$$

where t-statistic is reported in parentheses below each parameter estimate.

The diagonal elements in matrix Φ measure the impact of autoregressive terms for $\Delta MinRV_t$ and ΔILQ_t, which, as we expected, are statistically significant. The off-diagonal elements in matrix Φ provide insight on the cross-impacts between $\Delta MinRV_t$ and ΔILQ_t. We find the coefficient estimate measuring the impact of ΔILQ_{t-1} on $\Delta MinRV_t$ is 0.479 with a t-statistic of 4.50. This shows that a change in ILQ_{t-1} on the previous day has a positive and statistically significant impact on a change in $MinRV_t$ today. On the other hand, we do not find a statistically significant evidence that a change in ILQ_t today is driven by a change in $MinRV_{t-1}$ on the previous day; the coefficient estimate is -0.023 with the t-statistic of -1.47. As a robustness check, we re-estimate the VAR model by replacing $MinRV$ with RV in equation (13) and obtain the same conclusion. Further, looking at the coefficient estimates in Θ, we find that only the diagonal elements are statistically significant. This shows that the residual terms in the market spot variance L_t and the market illiquidity level ILQ_t do not affect each other. This result supports our modeling assumption in equations (9) and (10), for which the Brownian shocks dW_t^2 and dW_t^3 are independent.

Finally, we note that the variance dynamic that we consider in equation (9) falls under the class of two-factor stochastic volatility models, which have been shown to effectively explain the term structure of index option prices.[9] Our model differs from the existing two-factor volatility literature in that we allow the expected future variance to depend on the level of

[9]See for examples, Christoffersen, Heston, and Jacobs (2009), Egloff, Leippold, and Wu (2010), Bates (2012), and Andersen, Fusari, and Todorov (2015b).

spot variance, V_t, and spot market illiquidity, L_t, as shown below:

$$E_t[V_T] = \theta_V + (V_t - \theta_V)e^{-\kappa_V(T-t)} + [(L_t - \theta_L)\frac{\gamma\kappa_L}{\kappa_V - \kappa_L}](e^{-\kappa_V(T-t)} - e^{-\kappa_L(T-t)}). \tag{15}$$

This equation shows that the long-run mean of the spot variance is θ_V, and the mean-reversion speed to the long-run variance is denoted by κ_V. It also shows that the current level of market illiquidity positively affects the shape of the expected term structure of variance. Its impact, however, dissipates as the time horizon increases. This is seen from the third term on the right-hand side, which converges to 0 as time T goes to infinity. Without the term γL_t, the spot variance process reduces to the Heston (1993) model, and the expected future variance is given by the first two terms on the right-hand side of equation (15).

3.2 Benchmark Models

We consider two nested specifications of the SJVI model. In the first specification, we shut off illiquidity channel in the time-varying jump intensity dynamic by setting $\gamma_L = 0$ in equation (12). As the result, the probability of observing jumps depends on the level of spot variance and the latent state component as follows:

$$\text{SJV model:} \qquad \lambda_t = \Psi_t + \gamma_V V_t. \tag{16}$$

We refer to this as the stochastic jump intensity with variance (SJV) model. When Ψ_t is constant, it nests the affine jump intensity dynamic, $\lambda_t = \gamma_0 + \gamma_V V_t$, commonly adopted in time-varying jump studies (e.g., Pan, 2002; Bates, 2006).

The second nested specification that we study shuts off the impact of both market illiquidity and the spot variance on the jump probability. That is, we set γ_V and γ_L equal to zero in equation (12). This yields

$$\text{SJ model:} \qquad \lambda_t = \Psi_t, \tag{17}$$

We refer to this as the stochastic jump intensity model (SJ).

We keep all other aspects of the three models that we study identical. This allows us to focus solely on the role of market illiquidity and spot variance in determining the time-varying jump risk.

3.3 Filtering

Each of the three models we study contains three latent state variables: V_t, L_t, and Ψ_t. We extract the latent state variables using the square-root UKF of Van der Merwe and Wan

(2001). We apply the UKF method because the option prices data that we fit the models to are non-linear in the state variables.[10]

The state variables in the filtering equations evolve under the physical probability ($\mathbb{P}$) measure. We therefore need to define the state variables' dynamic under the physical measure. We do not impose risk premiums on the L_t and Ψ_t processes for simplicity and also because the literature has not yet provided clear guidance on how to model their risk premiums. As a result, there is no change to these two processes from $\mathbb{Q}$ to $\mathbb{P}$. We apply the commonly used functional form of the variance price of risk to the spot variance process, which is given by $\nu_V\sqrt{V_t}$ as in Heston (1993). This price of risk specification shifts the Brownian shock in equation (9) by $dW_t^{2,\mathbb{P}} = dW_t^2 - \nu_V\sqrt{V_t}dt$, where the superscript $\mathbb{P}$ denotes that it is evaluated under the physical probability measure. Applying this transformation, the resulting variance process under $\mathbb{P}$ can be written as

$$dV_t = \kappa_V^{\mathbb{P}}(\theta_V^{\mathbb{P}} - V_t)dt + \gamma dL_t + \xi_V\sqrt{V_t}dW_t^{2,\mathbb{P}}, \tag{18}$$

where we have the following parameter mappings $\kappa_V^{\mathbb{P}} = \kappa_V - \nu_V\xi_V$ and $\theta_V^{\mathbb{P}} = \theta_V\kappa_V/\kappa_V^{\mathbb{P}}$.

We discretize the $\mathbb{P}$-measure state dynamics using the conventional Euler scheme at the daily interval. The discretized state-space system can be written as follows:

$$V_{t+1} = V_t + \kappa_V^{\mathbb{P}}(\theta_V^{\mathbb{P}} - V_t)\Delta t + \gamma\kappa_L(\theta_L - L_t)\Delta t + \xi_V\sqrt{\Delta t V_t}\epsilon_{t+1}^1 + \gamma\xi_L\sqrt{\Delta t L_t}\epsilon_{t+1}^2 \tag{19}$$

$$L_{t+1} = L_t + \kappa_L(\theta_L - L_t)\Delta t + \xi_L\sqrt{\Delta t L_t}\epsilon_{t+1}^2 \tag{20}$$

$$\Psi_{t+1} = \Psi_t + \kappa_\Psi(\theta_\Psi - \Psi_t)\Delta t + \xi_\Psi\sqrt{\Delta t\Psi_t}\epsilon_{t+1}^3, \tag{21}$$

where the error terms ϵ_{t+1}^i, for $i = 1$ to 3, are i.i.d. standard normal. In the above state-space system, we set the time step $\Delta t = 1/252$ to reflect the daily discretization interval. To keep notation to a minimum, we apply the superscript $\mathbb{P}$ only to parameters under the physical measure that differ in values from their corresponding risk-neutral parameters.

We next describe the functional relationships linking the latent state variables to the observed data used in the estimation. The first observable is the illiquidity measure denoted by ILQ_t, which we introduced earlier in Section 2. The other observables that we use are daily at-the-money (ATM) and out-of-the-money (OTM) S&P 500 index options. These three sets of observables are used in the measurement equations in the UKF procedure. We write

[10] For recent papers using UKF as the filtering method, see Bakshi, Carr, and Wu (2008) and Filipović, Gourier, and Mancini (2016). We refer to Christoffersen, Dorion, Jacobs, and Karoui (2014) for technical details and comparison between different filtering methods.

the system of measure equations as follows:

$$\log(\mathrm{HLQ}_{t+1}) = \log(E_t[\int_t^{t+1} L_s ds]) + u_{t+1}^1 \quad (22)$$

$$ATM_{t+1}^O = ATM_{t+1}^M(V_{t+1}, L_{t+1}, \Psi_{t+1}) + u_{t+1}^2 \quad (23)$$

$$OTM_{t+1}^O = OTM_{t+1}^M(V_{t+1}, L_{t+1}, \Psi_{t+1}) + u_{t+1}^3, \quad (24)$$

where measurement errors u_{t+1}^i, for $i = 1$ to 3, are independent normal random variables with constant variances. The above filtering equations are applied to all trading days from January 2, 2004 to December 31, 2012, resulting in 2,262 observation days.

The latent spot illiquidity process in the state-space dynamic describes the instantaneous level of illiquidity at each moment and not at the aggregated daily level. To filter L_t from the daily observed market illiquidity measure, we integrate the spot illiquidity process over the day as shown in equation (22). Because the spot illiquidity measure is assumed to follow a square-root process, its daily integrated value is available in closed form. We use the log of effective spread instead in the measurement equation because the empirical distribution of effective spreads is close to log-normal.

Following Pan (2002), we collect two time series of closing mid-price of options quotes that we label ATM and OTM. We let ATM denote at-the-money call option that has moneyness, defined as the ratio of forward-to-strike price, closest to 1. Similarly, OTM refers to out-of-the-money put option that has moneyness closest to 0.95. For both ATM and OTM options, we retain contracts that have time-to-maturity closest to 30 calendar days. Figure 2 plots daily Black-Scholes option-implied volatilities calculated from the ATM and OTM contracts that we use in our study. As argued by Pan (2002), we use OTM options in the measurement equation as it provides the richest information on investors' expectation of crash probability in the stock market.

We follow Trolle and Schwartz (2009) and use Black-Scholes vega-weighted price as the functional form in the measurement equations for options fitting in equations (23)–(24). This method scales the value of options across time making their prices more comparable, which in turn facilitates the assumption of the normally distributed errors in the measure equations. Therefore, ATM_{t+1}^O and OTM_{t+1}^O in equations (23)–(24) represent the scaled ATM and OTM option prices observed at the end of day t. Similarly, the variables ATM_{t+1}^M and OTM_{t+1}^M denote the model-implied option prices scaled by their market Black-Scholes vega.

The models that we study fall within the affine jump-diffusion framework. Therefore, the conditional characteristic function of log stock price is available in exponential affine form. Following Duffie, Pan, and Singleton (2000), we derive the log affine functional form of the characteristic function in Appendix E. The coefficients in the characteristic function are

not all available in terms of elementary functions, thus, we solve for them numerically in the Ricatti system of equations. We use the fast Fourier transform (FFT) method first developed by Carr and Madan (1999) to numerically evaluate option prices.

Lastly, we note that at this stage, we do not need to specify the risk premiums associated with the first Brownian motion, dW_t^1, and the compound Poisson jumps, qdN_t, because they only alter the drift term of returns dynamics that is not part of the estimation. We discuss the specification of equity and jump risk premiums in a later section, where they are estimated using a time series of daily index returns.

3.4 Estimation

We estimate the models by maximizing the log-likelihood function resulting from the UKF step. We assume the measurement errors are conditionally normal, therefore, the time t conditional log-likelihood takes the following form:

$$l_t(\Theta) = -\frac{3}{2}\log(2\pi) - \frac{1}{2}\log(\det|\Omega_t|) - \frac{1}{2}(Y_t - Y_t)^T(\Omega_t)^{-1}(Y_t - Y_t), \tag{25}$$

where Y_t and Ω_t denote the ex ante forecasts of the mean and covariance matrix conditional on time $t-1$ information on observables Y_t. We let Θ denote the set of all parameters to be estimated.

In addition to the log-likelihood resulting from the measurement error equations, we follow Andersen, Fusari, and Todorov (2015a) and add a penalizing term that compares the filtered spot variance component, V_t, to the model-free estimate of spot variance calculated from high-frequency data. Incorporating this penalizing term, the conditional log-likelihood function that we estimate at time t is

$$L_t(\Theta) = l_t(\Theta) + \omega\log(\ (\sqrt{V_t^n} - \sqrt{V_t})^2\ ), \tag{26}$$

where $l_t(\Theta)$ is given in equation (25), V_t^n is the realized spot variance computed using one-minute grid returns from the S&P 500 index and V_t is the filtered spot variance from the UKF procedure. We describe the construction of the realized spot variance measure in more detail in Appendix F. Daily time-series dynamic of the realized spot volatility, $\sqrt{V_t^n}$, is shown in the bottom panel of Figure 2.

The tuning parameter ω in equation (26) is set equal to 0.05 following Andersen, Fusari, and Todorov (2015a).[11]The model parameters are then estimated by maximizing the sum of conditional log-likelihoods over the sample period from January 2, 2004 to December 31, 2012.

[11] As a robustness check, we verify that our main results remain virtually unchanged when picking different values of ω.

4 Results

4.1 Maximum likelihood estimates (MLEs)

Table 3 reports parameter estimates for the three models. The first, second, and third columns report results for the SJ, SJV and SJVI models, respectively. We report log-likelihood values of the three models in the bottom row.

We find that parameters governing the square-root dynamic of spot variance are well estimated. Their parameter estimates are fairly consistent across the models. The correlation estimates of the two Brownian shocks in return and spot variance, ρ, are about -35%, confirming the asymmetric return-variance relationship found in the extant literature. We find that the spot market illiquidity level, L_t, significantly impacts the level of spot variance, V_t. This is seen from the estimates of γ, which measure the contemporaneous relationship between market illiquidity and spot variance. We find that across the three models, the estimates γ are about 0.12. This suggests that a one-standard-deviation increase in the spot market illiquidity, L_t, would increase the spot variance level by about 12% after controlling for the persistence dynamic of the variance process.

The strong relationship we find between market illiquidity and return variance lends support to previous studies examining the relationship between return volatility and market trading activity. In particular, motivated by the mixture of distribution hypothesis (MDH), which assumes that volatility and volume simultaneously depend on a latent information process, past research efforts have been devoted to studying the relationship between stock return volatility and trading volume (e.g., Clark, 1973; Epps and Epps, 1976; Tauchen and Pitts, 1983). Nevertheless, the findings in this literature have been mixed and the understanding of relationships between information flows and trading activity has been an active research area. For instance, Lamoureux and Lastrapes (1990) estimate a GARCH volatility model and find that trading volume is the main driver of stock return volatility and that past stock return innovations became insignificant once trading volume is included in the model.[12] While we find that market illiquidity significantly drives the dynamic of spot variance, its effect does not eliminate the strong persistence in the variance dynamic. Further, the recent literature agrees that trading volume is an inadequate measure of market liquidity.[13] Given the recent availability of intraday trading data, we can more precisely measure market liquidity risk by calculating the cost of participating in the stock market (i.e., transaction cost). Our results estimated using a continuous-time model documenting a strong relationship between market

[12] In contrast, several studies find evidence conflicting with the MDH specification. These studies include Hiemstra and Jones (1994), Lamoureux and Lastrapes (1994), Richardson and Smith (1994) and Andersen (1996).

[13] See for examples, Lee, Mucklow, and Ready (1993), Jones (2002) and Fleming (2003).

illiquidity and return variance therefore contribute to this stream of literature.

Estimates of the jump-size mean, θ, and the jump-size standard deviation, δ, in Table 3 indicate that the jump dynamic that we estimate corresponds to crash risk in the stock market. The estimates of θ are negative and highly significant. The average jump mean size in daily index return is between -0.37% (SJV model) and -0.59% (SJ model). Therefore, the jump dynamic that we identify corresponds to large drops in daily S&P 500 index returns.

Table 3 shows that the SJ model has the largest magnitudes of θ and δ. This implies that crashes in the SJ model are larger and more dispersed in magnitude relative to the other two models. We next examine parameter estimates governing the time-varying jump intensity. First, we look at the dynamic of the latent jump-intensity-specific factor, Ψ_t. The magnitude of parameters driving the Ψ_t dynamic in the SJ model differs significantly from those in the other two models. For instance, the long-run mean θ_Ψ, the mean-reversion speed κ_Ψ, and the volatility ξ_Ψ of the jump-intensity-specific factor are significantly larger for the SJ model. These findings are expected because in the SJ model, jump intensity dynamic solely depends on the latent state variable Ψ_t. Further, these results confirm that the dynamic of jump intensity is time-varying and follows a mean-reverting process.

Table 3 shows that when we add covariates to the jump intensity dynamic in the SJV and SJVI models, the log-likelihood values of the model fit increases substantially. The improvement is large with an increase of about 5% relative to the SJ model. We therefore find strong support for modeling jump intensity as a function of economic covariates. Looking at the SJV model, we find the impact of spot variance on jump intensity, γ_V, is positive and statistically significant at the 5% level (t-statistic is 2.17). This finding is consistent with Pan (2002), Bates (2006), and Andersen, Fusari, and Todorov (2015b).

For the SJVI model, we find that when we add the market illiquidity measure to the jump intensity specification, the estimate of γ_V substantially decreases in magnitude and its statistically significance diminishes (t-statistic is 1.61). On the other hand, the impact of spot market illiquidity loads very strong (t-statistic is 13.87). This finding shows that the inclusion of market illiquidity as an economic covariate significantly weakens the relationship between jump intensity and spot variance. This finding is consistent with our conclusions from Table 1, which we obtained using regression analyses.

4.2 Time-Varying Volatility and Crash Risks

This section examines the time-series dynamics of market spot volatility and jump intensity. Table 4 reports descriptive statistics of daily jump intensity λ_t, spot variance V_t, and spot illiquidity L_t levels that we obtained using the UKF from 2004–2012. We find that the sample moments of daily spot illiquidity are almost identical across the three models. This

suggests that its dynamic is well identified when we extract their information from the daily market illiquidity measure ILQ_t calculated using effective bid-ask spreads.

Figure 3 plots the daily annualized jump intensity for the three models. The jump intensity dynamic of the SJ model is very volatile relative to the other two models. For instance, looking at the time-series statistics of λ_t in Table 4, we find the average expected number of jumps implied by this model is 1.19 per year, but with a median of 0.44 and a standard deviation of 5.95. This shows that the distribution of jump intensities filtered from the SJ model is highly skewed and dispersed. The average jump intensity implied by the SJ model is about half of the other models. However, the rarer nature of jumps observed in this model is compensated by its larger jump mean size of $\theta = -5.9\%$ per day as shown in Table 3.

Figure 3 shows the expected number of jumps in the SJ model increases dramatically during the 2008–2009 crisis period, while it is small outside the crisis period. The volatile nature of jump risk estimated from the SJ model explains why its filtered spot volatility dynamic is relatively smaller in magnitude (see Table 4) than in the other two models — because variations in index price dynamics are predominantly captured by jumps especially during the crisis period.

We find the jump intensity dynamic estimated from the SJV and SJVI models have comparable distributions with means of 2.5 and 2.9 jumps per year, respectively. In these two models, the levels of jump intensity are relatively stable before mid-2007, but rise after and peak in the fall of 2008. We believe the relatively more stable jump intensity dynamics observed in the SJV and SJVI models are due to improved identification resulting from the use of covariates in the jump intensity specification. This argument is supported by looking at the models' log-likelihood performance, which is substantially worse under the SJ model.

We next examine the economic contribution of the spot variance and market illiquidity to the jump intensity dynamic. Figure 4 plots the decomposition of daily jump intensity levels. Here, we decompose daily jump intensities filtered from the SJV model (top panel) and from the SJVI model (bottom panel) into their respective components.

For the SJV model, the top panel of Figure 4 shows that the market's spot variance is the main component driving jump intensity dynamic. We find that on average, 61% of the jump intensity level is explained by its covariation with the market's spot variance. The time-series average of its contribution is about 61%. We find the jump-intensity-specific factor Ψ_t explains about 39%. This finding shows that a non-trivially large portion of jump intensity cannot be explained by the dynamic of market's spot variance.

The bottom panel of Figure 4 shows the decomposition of daily jump intensities estimated from the SJVI model. Here, we find that the jump intensity dynamic is heavily dominated by its co-movement with equity market illiquidity. We plot daily percentage contributions of each jump intensity component in Figure 5. The results shown are largely consistent with the

findings in the bottom panel of Figure 4. We find that, on average, the market illiquidity factor explains about 64% of the jump probability in the SJVI model. In contrast to our findings for the SVJ model, we find the market's spot variance explains, on average, only 12%, with the remaining 24% contribution coming from the jump intensity factor, Ψ_t. Once we control for market illiquidity as an economic variable driving time-varying crash risk, the relative contribution of spot variance significantly diminishes.

The above findings offer important insight into the existing literature on index return models that has increasingly documented the importance of time-varying crash risk (e.g., Bates, 2006, 2012; Maheu, McCurdy, and Zhao, 2013). The common practice is to let jump intensity be an affine function of spot variance. This modeling approach is appealing because it is parsimonious. It identifies time-varying jump intensity as a constant multiple of the spot variance, thereby eliminating the need to introduce an additional state variable to the model. We find that our estimation results for the SVJ model provide some support for this modeling approach. However, we emphasize that the key economic variable that matters most from our results for modeling time-varying crash probability is not the market spot variance, but the market illiquidity factor. Lastly, our findings suggest the reason previous studies find a positive relationship between the stock market's time-varying crash risk and spot variance is because of their common exposure to market liquidity risk.

4.3 Impulse Response Function

We examine the impact of market illiquidity on the current and future crash probability using impulse response functions (IRFs). The question that the IRF answers is how much do the current and future values of crash intensity $\lambda_{t+\tau}$ respond to a one-standard-deviation increase in either the spot variance V_t, the spot market illiquidity L_t, or the level of latent state variable Ψ_t. The SJVI model that we propose yields an analytically tractable IRF for the jump intensity. For brevity, we report the IRF formula in Appendix G.

In Figure 6, we consider two dates where the levels of spot volatility are relatively high and low. The left-hand-side panels plot the IRF on March 11, 2009, with high spot volatility. The right-hand-side panels plot the IRF on January 8, 2004, which corresponds to the day with low spot volatility.

Looking at the IRF plots on the day with high volatility (left panels), we find that the impact of spot volatility dominates. A one-standard-deviation increase in the spot volatility V_t translates to an increase of 0.11 in jump intensity on the same day. Importantly, the impact of a shock to the spot volatility is very persistent with a half life of about two months. A one-standard-deviation increase in the market spot illiquidity L_t increases the jump intensity by about 0.07. Looking at the impact of a shock to the latent factor Ψ_t, we find that it is

trivially small at both the short- and long-run horizons.

We next look at the IRF plots on the day with low spot volatility. The state variables on this day are significantly less volatile and therefore the levels of IRF are much lower. On this low-volatility day, Figure 6 shows that a shock to market illiquidity dominates in term of magnitude as well as its lasting impact on the jump intensity. Similar to the day with high spot volatility, we find that the impact of a one-standard-deviation shock to the latent factor Ψ_t is small. Overall, Figure 6 shows that shocks to the spot volatility V_t and the market illiquidity L_t are highly persistent and drive most of the current and future increase in the probability of crash risk. However, the relative importance of V_t and L_t depends on the level of uncertainty in the market.

4.4 Forecast Error Variance Decomposition

We perform a forecast error variance decomposition (FEVD) on the jump intensity λ_τ dynamic for the SJVI model. This method helps determine the amount of information each variable contributes in explaining changes to the current and future crash probability.

The error from forecasting the jump intensity $\lambda_{t+\tau}$ with τ-period horizon conditional on day t is defined as

$$\epsilon_{\lambda,t+\tau} = \lambda_{t+\tau} - E_t[\lambda_{t+\tau}]. \tag{27}$$

The idea behind FEVD is to find how much of the variation in $\epsilon_{\lambda,t+\tau}$, i.e., $\mathrm{Var}_t[\epsilon_{\lambda,t+\tau}]$, can be explained by shocks to each state variable driving the jump intensity dynamic. In other words, the FEVD asks how much of the unexpected change in the jump intensity is explained by shocks to V_t, L_t, and Ψ_t. Derivation of the FEVD is algebraically tedious. For brevity, we report expressions in Appendix H.

In Figure 7, we plot the proportion of forecast error variance explained by innovations to the spot variance, V_t, and the spot market illiquidity, L_t, factors. The model parameters are obtained from their MLE estimates in Table 3, and their state variables are set equal to their long-run values. We do not plot the proportion of forecast error variance explained by Ψ_t because it is trivially small. The top two panels of Figure 7 plot the proportion of forecast error variance explained by V_t and L_t one day ahead, i.e., $\tau = 1$. These results provide insight on the source of information that most importantly impacts unexpected changes in crash probability at the very short horizon. We also plot the proportion of forecast error variance at a longer horizon, i.e., $\tau = 250$, which approximately corresponds to one year ahead. These results are shown in the bottom two panels.

We find that errors in the short-term forecast of the crash probability are mainly explained by shocks to the market illiquidity factor, with the exception of the crisis period when shocks

to the spot variance dominate. This finding is consistent with the results shown in Figure 5. In contrast, when we look at the sources of risk that explain errors in the long-term forecast of crash probability, we find that the market spot variance dominates. The bottom panels of Figure 7 show that the diffusive variance component in index returns contributes about 62% to unexpected changes in the long-run crash probability, while the market illiquidity component contributes around 38%. This finding suggests that changes in investors' perception about the long-run stock market crash risk is associated with the market uncertainty level. On the other hand, unexpected changes in crash probability in the near horizon are mostly explained by shocks to market illiquidity.

4.5 Option Fit

We also compare the three models based on their in-sample option fits. We define in-sample option pricing error as the vega-weighted root mean squared error (VWRMSE) in fitting observed Black-Scholes vega-weighted option prices obtained from the UKF procedure

$$\text{VWRMSE(ATM)} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(ATM_{t+1}^{O} - ATM_{t+1}^{M})^2}, \tag{28}$$

where ATM_{t+1}^{M} denotes the ex ante forecast of vega-weighted ATM option price at time $t+1$. Option pricing error for OTM options are computed in a similar way.

Table 5 reports in-sample option pricing errors for the three models. We separate the sample into three sub-periods of three years each. Overall pricing errors are very similar in magnitude between SJ and SJV models, where the SJV model performs better during the crisis period while the SJ model has lower pricing errors during normal times. More importantly, the SJVI model has a superior in-sample option fit in most of the periods for both ATM and OTM options. Thus, the improvement in fitting OTM options using the SJVI model suggests that its jump intensity specification is better-suited for capturing the jump intensity dynamic embedded in the index options.

4.6 Risk Premiums

Using the risk-neutral parameter estimates in Table 3 and daily filtered states variables $\{\hat{V}_t, \hat{L}_t, \hat{\Psi}_t\}$ estimated previously, we infer the risk premium parameters. This is done by estimating the model on daily S&P 500 index returns from 2004–2012, and keeping the parameters that are not affected by the change of probability measures fixed. This approach to identify risk premiums was also employed in Andersen, Fusari, and Todorov (2015b) and Christoffersen, Fournier, and Jacobs (2015).

We assume the conventional form of the pricing kernel that preserves the affine structure of the model under the physical measure. The prices of risk associated with the four Brownian motions are given by

$$dW_t^{1,\mathbb{P}} = dW_t^1 - \nu_1\sqrt{V_t}dt \tag{29}$$

$$dW_t^{2,\mathbb{P}} = dW_t^2 - \nu_V\sqrt{V_t}dt \tag{30}$$

$$dW_t^{3,\mathbb{P}} = dW_t^3 \tag{31}$$

$$dW_t^{4,\mathbb{P}} = dW_t^4 \tag{32}$$

The parameter ν_1 in equation (29) corresponds to the price of risk parameter for the first Brownian innovation in the return process. Recall that ν_V is the price of risk parameter for the volatility innovation that we estimated from options and realized spot variance as part of the UKF. Its estimate is reported in Table 3. Recall also that we do not impose any risk premium assumptions on the third and fourth Brownian motions corresponding to the liquidity and latent jump intensity innovations, respectively.

We follow Pan (2002) and assume the difference between jump distributions under the physical and risk-neutral measures derives from the jump-size risk premium, ν_θ, defined as the difference between jump-size means, $\theta^P - \theta$. The dynamic of log-stock price under the physical probability measure can be written as

$$\begin{aligned} d\log(S_t) = {} & (r - \frac{1}{2}V_t - \xi^P\lambda_t + (\sqrt{1-\rho^2}\nu_1 + \rho\nu_v)V_t)dt + \\ & \sqrt{V_t}(\sqrt{1-\rho^2}dW_t^{1,\mathbb{P}} + \rho dW_t^{2,\mathbb{P}}) + q_t dN_t^{\mathbb{P}} \end{aligned} \tag{33}$$

where $\xi^P = \exp(\theta^P + \frac{1}{2}\delta^2)$ is the jump compensator under the physical measure. Comparing the $\mathbb{P}$-measure return dynamic in equation (33) to the $\mathbb{Q}$-measure return dynamic in equation (8) shows that the equity risk premium, π_t, can be written as

$$\pi_t = (\xi^P - \xi)\lambda_t + (\sqrt{1-\rho^2}\nu_1 + \rho\nu_V)V_t \tag{34}$$

$$= (\xi^P - \xi)\lambda_t + \nu_S V_t, \tag{35}$$

where we define $\nu_S = \sqrt{1-\rho^2}\nu_1 + \rho\nu_v$ in equation (35).

Using the filtered state variables, $\{\hat{V}_t, \hat{L}_t, \hat{\Psi}_t\}$, we apply daily discretization to the return process and estimate the risk premium parameters ν_θ and ν_S using MLE while fixing all other parameters. The estimate for ν_1 is then inferred from ν_S. Appendix F shows the discretization of the continuous-time model, and presents the log-likelihood function for fitting the return process.

Table 6 reports estimation results of the risk premium parameters. We find that the

jump risk premium parameter ν_ϑ is well identified in all models. The estimates for ν_ϑ are statistically significant with the confidence level of 1% or greater. On the other hand, estimates of the diffusive risk premium parameter ν_S are only marginally significant. These findings are consistent with Pan (2002) who finds that the jump risk premium is more easily identified from index option prices, while risk premiums associated with the diffusive and variance risks are more difficult to estimate. Table 6 also reports estimates for the price of risk coefficient ν_1 associated with the first Brownian motion. Their values are inferred from the corresponding estimates of ν_S in Table 6, and ν_V in Table 3. Because ν_1 is indirectly inferred, we do not report its t-statistic. This parameter can be usefully thought of as the price risk for exposure to the diffusive component in index return.

Using the estimates reported in Tables 6 and 3, we quantify the economic magnitude of each risk premium component in terms of annualized excess returns. Equation (35) shows that the equity risk premium can be decomposed into two main components. The first component represents the compensation for bearing stock market crash risk, $(\xi^P - \xi)\lambda_t$. The second component represents the compensation for bearing the stock market's diffusive return and variance risks, $\nu_S V_t$. For brevity, we refer to $\nu_S V_t$ as diffusive risk in the equity risk premium.

We first look at the compensation for bearing stock market crash risk. For each model, we calculate the long-run jump risk premium level $(\xi^P - \xi)\lambda_t$, where λ_t is the annualized time-series mean of the jump intensity dynamic reported in Table 4. We find the compensation for bearing the market's crash risk for the SJ, SJV, and SJVI models are 3.0%, 5.4%, and 4.9% in annualized excess returns, respectively. The jump risk premium estimate implied by the SJ model is lower than the other two models. This finding reflects the relatively lower jump intensity levels that we find for the SJ model. The jump risk premium estimates implied by the SJV and SJVI models are mostly consistent with prior studies that estimate a time-varying jump risk model on the S&P 500 index over a similar sample period. For instance, Ornthanalai (2014) estimates the jump risk premium implied by the compound Poisson jump process over the 1996–2012 period and finds that its magnitude is 4.5% per year. Using index options and returns data from an earlier time period, i.e., 1989–1996, Pan (2002) finds that the implied jump risk premium is 3.5% per annum.

We next look at the compensation for bearing stock market diffusive risk. This is calculated as $\nu_S V_t$, where V_t is the time-series mean of the annualized variance reported in Table 4. We find the compensation for bearing diffusive risk for the SJ, SJV, and SJVI models is 7.48%, 4.53%, and 3.54% in annualized excess returns, respectively. The relatively larger magnitude of diffusive risk premium found in the SJ model is expected. This is because the SJ model has the lowest jump risk premium level and hence it must rely on the diffusive risk premium component to match the level of equity risk premium found in the data.

The realized equity premium calculated using daily index returns data over the 2004–2012

period is 8.7% per year. The total equity premiums that we find for the SJ, SJV, and SJVI models are 10.48%, 9.95%, and 8.45%, in annualized terms, respectively. Our estimates of the total equity premium are therefore consistent with the value calculated using daily returns data. This finding suggests that the magnitudes of equity risk premium implied by our models are economically plausible.

5 Robustness

5.1 Circuit Breakers

Following the flash crash incident on May 6, 2010, the Securities and Exchange Commission (SEC) installed "circuit breakers" on 404 NYSE-listed S&P 500 stocks on June 16, 2010 to halt trading for five minutes if any stock experiences more than 10% movement, either up or down, in a five-minute period. This new trading rule potentially affects our aggregate illiquidity measure constructed from individual firms' effective spreads, thereby altering the impact of market illiquidity on jump probability. We test whether this change in market-trading rules alters our findings on the influence of market illiquidity on time-varying crash risk.

We take June 16, 2010 as the date of exogenous shift in the market-trading structure. Specifically, we divide our sample into two periods, one starting on January 4, 2004 and ending on June 15, 2010, and the other starting on June 16, 2010 and ending on December 31, 2012. We use the same jump intensity specification as in the SJVI model for both subsamples, as below. The model parameters are estimated separately, yielding two sets of parameter estimates. We summarize the results below. For brevity, we report only coefficient loadings on the covariates in the jump intensity dynamic. The t-statistic for each parameter is reported in parentheses underneath its estimate.[14]

$$\text{Before the circuit breaker:} \qquad \lambda_t = \Psi_t + \underset{(1.71)^*}{19.69\, V_t} + \underset{(9.04)^{***}}{9.24\, L_t} \tag{36}$$

$$\text{After the circuit breaker:} \qquad \lambda_t = \Psi_t + \underset{(1.47)}{18.47\, V_t} + \underset{(1.98)^{**}}{7.98\, L_t}. \tag{37}$$

We find that the loading coefficient on the market illiquidity factor is smaller in latter period, being 7.98, relative to the estimate of 9.24 in the earlier period. This suggests that the introduction of circuit breakers has slightly reduced the impact of market illiquidity on jump intensity, perhaps, by eliminating sudden contiguous large movements in equity prices that were often identified as symptoms of liquidity shortage. Nevertheless, equations (36)–(37) show

[14]***, **, and * denote statistical significance at the 1, 5, and 10 confidence levels, respectively.

that the implementation of circuit breakers does not materially impact the importance of the market illiquidity channel. Both coefficients are statistically significant at 1% and 5% levels, respectively, with a lower t-statistic for the latter period because of its much smaller sample size. The loading coefficients on market spot variance V_t for the pre- and post-circuit breaker periods are 19.69 and 18.47, respectively. These magnitudes are similar to the estimate of 18.38 found using the full sample period. In both periods, the impact of the coefficient estimates on V_t are statistically weak. These results confirm the robustness of our estimates in Table 3.

5.2 Alternative Iliquidity Measure

We have so far defined market illiquidity using the aggregate *relative* effective spread of S&P 500 constituents. This measure captures the aggregate transaction cost of participating in the stock market and has been shown in Aït-Sahalia and Yu (2009), and Goyenko, Holden, and Trzcinka (2009) to be a good proxy for market illiquidity. This section tests whether our results are robust to other illiquidity measures.

First, we use dollar effective spread as an alternative measure for the market-trading cost. It is calculated as the absolute dollar difference between the transaction price and the prevailing mid-price of each transaction instead of the relative percentage to the mid-price as before. More precisely, the dollar effective spread measure associated with each transaction k on day t for firm i is defined as

$$\$ILQ^i_{t,k} = 2|S^{i,P}_{t,k} - S^{i,M}_{t,k}|. \tag{38}$$

We aggregate $\$ILQ^i_{t,k}$ across S&P 500 index constituents to construct the daily measure of market illiquidity.

Our second alternative market illiquidity measure is from Amihud (2002). On each day t, we compute the Amihud illiquidity measure for each firm i as a fraction of absolute return, $|r_i|$, over dollar trading volume, $DVol_{i,t}$, that day:

$$ALIQ^i_t = \sum_{i=1}^{N} \frac{|r_{i,t}|}{DVol_{i,t}}. \tag{39}$$

The daily Amihud market illiquidity measure for the stock market is then calculated as an equal-weighted average of individual firms' Amihud illiquidity measure S&P 500 constituents.

The middle and bottom panels of Figure A.2 plot the time-series dynamic of the two alternative market illiquidity measures. We plot the relative effective spread in the top panel. All illiquidity measures are normalized to have the same sample mean. This normalization method does not impact our results because the absolute level does not matter for our specification.

Figure A.2 shows that the dollar effective spread measure is similar to the relative effective spread measure, although with some small differences during the crisis period. The Amihud illiquidity measure is much noisier than the other two measures calculated from intraday bid-ask spreads.

We re-estimate the SJVI model using the market dollar effective spread and the Amihud illiquidity measure. We report the results in Table 7. Overall parameter estimates are fairly consistent compared with those estimated using relative effective spreads shown in Table 3. We find the coefficient estimate γ_V is similar in magnitude and is not statistically significant in either alternative illiquidity measure that we use. Meanwhile, the coefficient estimate for γ_L is similar for the Amihud illiquidity measure, and even larger for the dollar effective spread measure. In all cases, γ_L remains statistically significant. In sample option pricing errors are higher using the Amihud illiquidity measure, as expected from the noisiness of the measure. On the other hand, the in-sample fit of options shows very marginal improvement using the dollar effective spread. To save space, the option pricing results are not reported here. Overall, our main conclusions remain qualitatively unchanged. We conclude that our main results are robust to different definitions of market illiquidity.

6 Conclusion

We study the role of market illiquidity in explaining the time-varying market crash risk in the S&P 500 index. We estimate a continuous-time model with stochastic volatility and crash probability. We introduce market liquidity as an observable variable to the model by allowing it to affect the dynamics of spot variance and jump risk intensity. We follow the recent empirical literature in market illiquidity risk and measure the daily stock market illiquidity level using volume-weighted intraday bid-ask spreads of all securities constituting the S&P500 index. We estimate the model over 2004–2012 using daily S&P 500 index options, realized spot variance and market illiquidity measure, and find that 64% of time-varying crash risk is due to the stock market's exposure to market illiquidity. The influence of market illiquidity dominates other factors that we examined, including the market's spot variance. This is with an exception of the 2008 crisis, when the influence of spot variance dominates and the contribution of market illiquidity falls to about 30%. Overall, our paper highlights the importance of market-trading frictions in index return models and suggests that the time-varying crash risk mostly reflects investors' fear of market illiquidity.

References

Aït-Sahalia, Y., and J. Yu, 2009, "High Frequency Market Microstructure Noise Estimates and Liquidity Measures," *Annals of Applied Statistics*, 3(1), 422–457.

Amaya, D., P. Christoffersen, K. Jacobs, and A. Vasquez, 2015, "Does Realized Skewness Predict the Cross-section of Equity Returns?" *Journal of Financial Economics*, 118(1), 135–167.

Amihud, Y., 2002, "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects," *Journal of Financial Markets*, 5(1), 31–56.

Andersen, T. G., 1996, "Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility," *The Journal of Finance*, 51(1), 169–204.

Andersen, T. G., L. Benzoni, and J. Lund, 2002, "An Empirical Investigation of Continuous-Time Equity Return Models," *Journal of Finance*, 57, 1239–1284.

Andersen, T. G., T. Bollerslev, and F. Diebold, 2007, "Roughing it Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility," *Review of Economics and Statistics*, 89(4), 701–720.

Andersen, T. G., D. Dobrev, and E. Schaumburg, 2012, "Jump-Robust Volatility Estimation Using Nearest Neighbor Truncation," *Journal of Econometrics*, 169(1), 75–93.

Andersen, T. G., N. Fusari, and V. Todorov, 2015a, "Parametric Inference and Dynamic State Recovery From Option Panels," *Econometrica*, 83(3), 1081–1145.

———, 2015b, "The Risk Premia Embedded in Index Options," *Journal of Financial Economics*, 117(3), 558–584.

Bakshi, G., P. Carr, and L. Wu, 2008, "Stochastic Risk Premiums, Stochastic Skewness in Currency Options, and Stochastic Discount Factors in International Economics," *Journal of Financial Economics*, 87(1), 132–156.

Bakshi, G., and D. Madan, 2000, "Spanning and Derivative Security Valuation," *Journal of Financial Economics*, 55(2), 205–238.

Barndorff-Nielsen, O. E., and N. Shephard, 2004, "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2(1), 1–37.

Bates, D., 2006, "Maximum Likelihood Estimation of Latent Affine Processes," *Review of Financial Studies*, 26(9), 909–965.

Bates, D., 2012, "U.S. Stock Market Crash Risk, 1926-2010," *Journal of Financial Economics*, 105(2), 229–259.

Bradley, D., J. Clarke, S. Lee, and C. Ornthanalai, 2014, "Are Analysts' Recommendations Informative? Intraday Evidence on the Impact of Time Stamp Delays" *The Journal of Finance*, 69(2), 645–673.

Broadie, M., M. Chernov, and M. Johannes, 2007, "Model Specification and Risk Premia: Evidence from Futures Options," *Journal of Finance*, 63, 1453–1490.

Carr, P., and D. B. Madan, 1999, "Option Valuation Using the Fast Fourier Transform," *Journal of Computational Finance*, 2(4), 61–73.

Carr, P., and L. Wu, 2003, "What Type of Process Underlies Options? A Simple Robust Test" *The Journal of Finance*, 58(6), 2581–2610.

Cespa, G., and T. Foucault, 2014, "Illiquidity Contagion and Liquidity Crashes," *Review of Financial Studies*, 27(6), 1615–1660.

Chae, J., 2005, "Trading Volume, Information Asymmetry, and Timing Information," *The Journal of Finance*, 60(1), 413–442.

Chang, B. Y., P. Christoffersen, and K. Jacobs, 2013, "Market Skewness Risk and the Cross Section of Stock Returns," *Journal of Financial Economics*, 107(1), 46–68.

Chen, J., H. Hong, and J. C. Stein, 2001, "Forecasting Crashes: Trading Volume, Past Returns, and Conditional Skewness in Stock Prices," *Journal of Financial Economics*, 61(3), 345–381.

Christoffersen, P., C. Dorion, K. Jacobs, and L. Karoui, 2014, "Nonlinear Kalman Filtering in Affine Term Structure Models," *Management Science*, 60(9), 2248–2268.

Christoffersen, P., M. Fournier, and K. Jacobs, 2015, "The Factor Structure in Equity Options," *Rotman School of Management Working Paper*.

Christoffersen, P., S. Heston, and K. Jacobs, 2009, "The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work So Well," *Management Science*, 55(12), 1914–1932.

Christoffersen, P., K. Jacobs, and C. Ornthanalai, 2012, "Dynamic Jump Intensities and Risk Premiums: Evidence from S&P 500 Returns and Options," *Journal of Financial Economics*, 106(3), 447–472.

Chung, K. H., and C. Chuwonganant, 2014, "Uncertainty, Market Structure, and Liquidity," *Journal of Financial Economics*, 113(3), 476–499.

Clark, P. K., 1973, "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41(1), 135–155.

Duffie, D., J. Pan, and K. Singleton, 2000, "Transform Analysis and Asset Pricing for Affine Jump-diffusions," *Econometrica*, 68(6), 1343–1376.

Egloff, D., M. Leippold, and L. Wu, 2010, "The Term Structure of Variance Swap Rates and Optimal Variance Swap Investments," *Journal of Financial and Quantitative Analysis*, 45, 1279–1310.

Epps, T. W., and M. L. Epps, 1976, "The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis," *Econometrica*, 44(2), 305–321.

Eraker, B., 2004, "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices" *Journal of Finance*, 59(3), 1367–1404.

Filipović, D., E. Gourier, and L. Mancini, 2016, "Quadratic Variance Swap Models," *Journal of Financial Economics*, 119(1), 44–68.

Fleming, M. J., 2003, "Measuring Treasury Market Liquidity," *Economic Policy Review*, (9), 83–108.

Gennotte, G., and H. Leland, 1990, "Market Liquidity, Hedging, and Crashes," *The American Economic Review*, 80(5), 999–1021.

Goyenko, R. Y., C. W. Holden, and C. A. Trzcinka, 2009, "Do Liquidity Measures Measure Liquidity?" *Journal of Financial Economics*, 92(2), 153–181.

Heston, S., 1993, "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6(2), 327–343.

Hiemstra, C., and J. D. Jones, 1994, "Testing for Linear and Nonlinear Granger Causality in the Stock Price-Volume Relation," *The Journal of Finance*, 49(5), 1639–1664.

Huang, J., and J. Wang, 2009, "Liquidity and Market Crashes," *Review of Financial Studies*, 22(7), 2607–2643.

Huang, X., and G. Tauchen, 2005, "The Relative Contribution of Jumps to Total Price Variance," *Journal of Financial Econometrics*, 3(4), 456–499.

Jiang, G. J., and T. Yao, 2013, "Stock Price Jumps and Cross-Sectional Return Predictability," *Journal of Financial and Quantitative Analysis*, 48, 1519–1544.

Jones, C. M., 2002, "A Century of Stock Market Liquidity and Trading Costs," *Working Paper*.

Kozhan, R., A. Neuberger, and P. Schneider, 2014, "The Skew Risk Premium in the Equity Index Market," *Review of Financial Studies*, 26(9), 2174–2203.

Lamoureux, C. G., and W. D. Lastrapes, 1990, "Heteroskedasticity in Stock Return Data: Volume Versus GARCH Effects," *Journal of Finance*, 45(1), 221–229.

———, 1994, "Endogenous Trading Volume and Momentum in Stock-Return Volatility," *Journal of Business & Economic Statistics*, 12(2), 253–260.

Lee, C. M. C., B. Mucklow, and M. J. Ready, 1993, "Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis," *Review of Financial Studies*, 6(2), 345–374.

Maheu, J., and T. McCurdy, 2004, "News Arrival, Jump Dynamics, and Volatility Components for Individual Stock Returns," *Journal of Finance*, 59(2), 755–793.

Maheu, J. M., T. H. McCurdy, and X. Zhao, 2013, "Do Jumps Contribute to the Dynamics of the Equity Premium?" *Journal of Financial Economics*, 110(2), 457–477.

Mancini, C., 2009, "Non-Parametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps," *Scandinavian Journal of Statistics*, 36(2), 270–296.

Ornthanalai, C., 2014, "Lévy Jump Risk: Evidence from Options and Returns," *Journal of Financial Economics*, 112(1), 69–90.

Pan, J., 2002, "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study," *Journal of Financial Economics*, 63, 3–50.

Richardson, M., and T. Smith, 1994, "A Direct Test of the Mixture of Distributions Hypothesis: Measuring the Daily Flow of Information," *Journal of Financial and Quantitative Analysis*, 29, 101–116.

Santa-Clara, P., and S. Yan, 2010, "Crashes, Volatility, and the Equity Premium: Lessons from S&P 500 Options," *Review of Economics and Statistics*, 92(2), 435–451.

Tauchen, G., and M. Pitts, 1983, "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica*, 51(2), 485–505.

Trolle, A. B., and E. S. Schwartz, 2009, "Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives," *Review of Financial Studies*, 22(11), 4423–4461.

Van der Merwe, R., and E. A. Wan, 2001, "The Square-Root Unscented Kalman Filter for State and Parameter-Estimation," in *Proceedings. 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 6, pp. 3461–3464.

White, H., 1980, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48(4), 817–838.

Table 1: Regression Model on Realized Jump Variation (RJV)

	Realized Jump Variation Next Day: RJV_{t+1}					
	(1)	(2)	(3)	(4)	(5)	(6)
$MinRV_t$	0.071***			0.038		0.037
	(2.63)			(1.36)		(1.31)
ILQ_t		0.091***		0.061***	0.070**	0.045**
		(3.97)		(4.25)	(2.13)	(2.01)
VIX_t			0.046***		0.014	0.011
			(4.50)		(1.19)	(0.94)
Ret_t	-0.058	-0.057	-0.038	-0.054	-0.050	-0.049
	(-0.97)	(-0.94)	(-0.63)	(-0.92)	(-0.79)	(-0.79)
RJV_t	-0.196	-0.075	-0.006	-0.185	-0.068	-0.177
	(-1.19)	(-0.94)	(-0.09)	(-1.22)	(-0.85)	(-1.15)
Adjusted R^2	29.3%	30.5%	29.0%	32.1%	30.7%	32.2%

Notes: We report estimated coefficients and t-statistics from the predictive OLS regression on the non-parametrically estimated variance component in daily S&P 500 index returns that is due to jumps. The sample period is from January 2, 2004 to December 31, 2012. The dependent variable is the realized jump variation (RJV) calculated using high-frequency intraday trades on S&P 500 cash index. The independent variables include lagged realized variance estimator, $MinRV$, from Andersen, Dobrev, and Schaumburg (2012), which measures variations in daily S&P 500 index returns that are associated with non-jump risk; market illiquidity proxy, ILQ, measured by daily averaged effective spreads across firms in the S&P 500 constituents; option-implied volatility index, VIX, obtained from the CBOE; and log return of S&P 500 index. All variables are lagged by one day. RJV, $MinRV$, ES, and VIX are expressed in annualized terms by multiplying their daily measure by 252. We also include the autoregressive term for RJV in the regression. Year and day-of-the week fixed effects are included. Coefficient estimates on the fixed-effect terms are not reported here to save space. We report heteroskedasticity-consistent t-statistic in parenthesis "( )" below each parameter estimate. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively, based on the heterosedasticity-consistent t-statistic. Adjusted R-squared for each regression model is reported in the bottom row.

Table 2: Regression Model on Change in Risk-Neutral Skewness

	Change in Risk-Neutral Skewness: $\Delta RNSkew_t$			
	(1)	(2)	(3)	(4)
$\Delta MinRV_t$		-0.005		0.450*
		(-0.03)		(1.69)
ΔILQ_t			-0.622*	-1.144**
			(-1.79)	(-2.48)
$Return_t$	-0.549***	-0.551***	-0.826***	-0.857***
	(-3.28)	(-3.00)	(-3.40)	(-3.70)
$\Delta RNSkew_{t-1}$	0.455***	0.455***	0.449***	0.451***
	(18.93)	(18.86)	(18.59)	(18.74)
AICC	-2.181	-2.180	-2.182	-2.182
R^2	23.4%	23.4%	23.5%	23.7%

Notes: We report regression results on daily changes in risk-neutral skewness of S&P 500 index returns, $\Delta RSkew_t$. The sample period is from January 2, 2004 to December 31, 2012. Risk-neutral skewness, $RSkew_t$, on day t is calculated from end-of-the-day S&P 500 index option prices with maturity closest to 30 days. We use the nonparametric method of Bakshi and Madan (2000) to calculate the 30-day forward-looking risk-neutral skewness measure. The independent variables include change in realized variance estimator, $\Delta MinRV$, from Andersen, Dobrev, and Schaumburg (2012); change in market illiquidity proxy; ΔILQ, measured by daily averaged effective spreads across firms in the S&P 500 constituents; *Return*, log S&P 500 return. Each specification is estimated using maximum likelihood. We use an ARMA(2,1) structure in the regression model, which is determined by the LjungBox test to sufficiently remove cross-correlations in the residuals. We control for seasonality due to the day-of-the-week effect. We report coefficient estimates on the two autoregressive terms. For brevity, we do not report coefficient estimates on the moving-average error term and day-of-the-week fixed effects. Robust t-statistic is reported in parenthesis below each parameter estimate. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. The last row reports regression diagnostics based on the Akaike information criterion (AICC) and R^2 metric.

Table 3: Maximum Likelihood Estimates: 2004–2012

	(1) SJ Model $\lambda_t = \Psi_t$	(2) SJV Model $\lambda_t = \Psi_t + \gamma_V V_t$	(3) SJVI Model $\lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t$
Parameter	Estimate	Estimate	Estimate
Panel A. Filtered state dynamics			
κ_V	3.4219	3.5652	3.5490
	(7.75)	(8.39)	(4.73)
θ_V	0.0307	0.0309	0.0312
	(8.86)	(11.00)	(5.66)
ξ_V	0.3361	0.3425	0.3456
	(99.50)	(22.97)	(42.49)
ν_V	1.6040	1.5592	1.5536
	(1.06)	(0.42)	(0.71)
κ_L	2.4159	2.3445	2.3532
	(3.93)	(1.27)	(4.61)
θ_L	0.1775	0.1822	0.1709
	(7.64)	(5.16)	(6.69)
ξ_L	0.1490	0.1510	0.1583
	(37.50)	(6.10)	(36.80)
κ_Ψ	0.9725	0.6606	0.6616
	(5.55)	(0.83)	(2.25)
θ_Ψ	1.6186	0.1022	0.1014
	(5.00)	(3.34)	(1.66)
ξ_Ψ	0.4015	0.2036	0.2037
	(2.47)	(7.22)	(1.79)
ρ	-0.3434	-0.3511	-0.3527
	(19.03)	(2.90)	(5.92)
Panel B. Jump-size parameters			
θ	-0.0590	-0.0367	-0.0373
	(34.98)	(10.41)	(26.28)
δ	0.0470	0.0318	0.0314
	(55.16)	(13.48)	(29.84)
Panel C. Loadings on covariates			
γ	0.1201	0.1167	0.1180
	(5.22)	(3.52)	(7.95)
γ_V		52.7228	18.3801
		(2.17)	(1.61)
γ_L			9.2590
			(13.87)
Log-Likelihood:	7,307.32	7,641.24	7,707.29

Notes: We report MLEs of the three time-varying jump models: SJ, SJV, and SJVI. The sample period is from January 2, 2004 to December 31, 2012. Each model is estimated using daily OTM and ATM S&P 500 index options, averaged effective spreads of S&P 500 constituents, and spot variance estimated from one-minute high-frequency S&P 500 futures data. We maximize the log likelihood function in equation (26). The state variables are estimated using the UKF. We report t-statistic calculated using the outer product of the gradient in parenthesis below each parameter estimate.

Table 4: Descriptive Statistics of Filtered Jump Intensities and Spot Variances

	(1) SJ Model $\lambda_t = \Psi_t$	(2) SJV Model $\lambda_t = \Psi_t + \gamma_V V_t$	(3) SJVI Model $\lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t$
Panel A. Jump intensity λ_t			
Mean	1.1922	2.5273	2.9172
Median	0.4367	1.6013	2.3124
Std. Dev.	5.9503	3.3500	2.1723
25 percentile	0.1818	0.9441	1.8025
75 percentile	0.8955	2.7791	3.2946
Panel B. Spot variance V_t			
Mean	0.0258	0.0290	0.0316
Median	0.0159	0.0156	0.0146
Std. Dev.	0.0334	0.0429	0.0561
25 percentile	0.0099	0.0095	0.0086
75 percentile	0.0280	0.0297	0.0273
Panel C. Spot illiquidity L_t			
Mean	0.1681	0.1681	0.1681
Median	0.1539	0.1539	0.1538
Std. Dev.	0.0558	0.0557	0.0558
25 percentile	0.1374	0.1375	0.1374
75 percentile	0.1741	0.1742	0.1742

Notes: We report the descriptive statistics of filtered jump intensities λ_t, spot variances V_t, and spot illiquidity L_t for three models: SJ, SJV, and SJVI. The variables are reported in annualized terms by multiplying their daily values by 252. We obtain the filtered state variables from the UKF step in the MLE estimation. Parameter estimates of the three models are reported in Table 3.

Table 5: VWRMSE of Different Models

	(1) SJ Model $\lambda_t = \Psi_t$	(2) SJV Model $\lambda_t = \Psi_t + \gamma_V V_t$	(3) SJVI Model $\lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t$
Panel A. VWRMSE by sub-period for OTM options			
2004–2006	3.21%	3.38%	3.21%
2007–2009	10.41%	9.99%	8.72%
2010–2012	6.57%	6.57%	6.00%
2004–2012	7.34%	7.17%	6.39%
Panel B. VWRMSE by sub-period for ATM options			
2004–2006	1.12%	1.68%	1.21%
2007–2009	7.69%	7.58%	5.82%
2010–2012	4.50%	4.70%	4.00%
2004–2012	5.18%	5.24%	4.14%

Notes: We report in-sample fit for the three models: SJ, SJV, and SJVI. The models are estimated using MLE. Equation (26) shows the log likelihood function. Panel A reports in-sample option pricing errors for OTM and panel B reports the pricing errors for ATM options. Option pricing errors are obtained from the measurement equations in the UKF step. The numbers reported are vega-weighted root mean squared error (VWRMSE). For ATM options, the VWRMSE is calculated as

$$\text{VWRMSE(ATM)} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(ATM_{t+1}^{O} - ATM_{t+1}^{M}\right)^2}$$

where ATM_{t+1}^{M} denotes the ex ante forecast of vega-weighted ATM option price at time $t+1$, and ATM_{t+1}^{O} denotes the vega-weighted ATM option price observed in the data. We use option vega reported in the Ivey Optionmetrics database to scale option prices, which makes their levels more comparable across moneyness and time (Trolle and Schwartz, 2009). The VWRMSE for OTM options is calculated similarly.

Table 6: Risk Premium Parameters Estimated from Daily Returns: 2004–2012

	(1) SJ Model $\lambda_t = \Psi_t$	(2) SJV Model $\lambda_t = \Psi_t + \gamma_V V_t$	(3) SJVI Model $\lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t$
Parameter	Estimate	Estimate	Estimate
$\nu_s = \sqrt{1-\rho^2}\nu_1 + \rho\nu_v$	2.900	1.562	1.121
	(1.89)	(1.39)	(1.54)
ν_θ	0.048	0.025	0.019
	(11.76)	(4.23)	(2.24)
ν_1	3.674	2.253	1.784
Log-Likelihood:	7,125.88	7,190.94	7,204.12

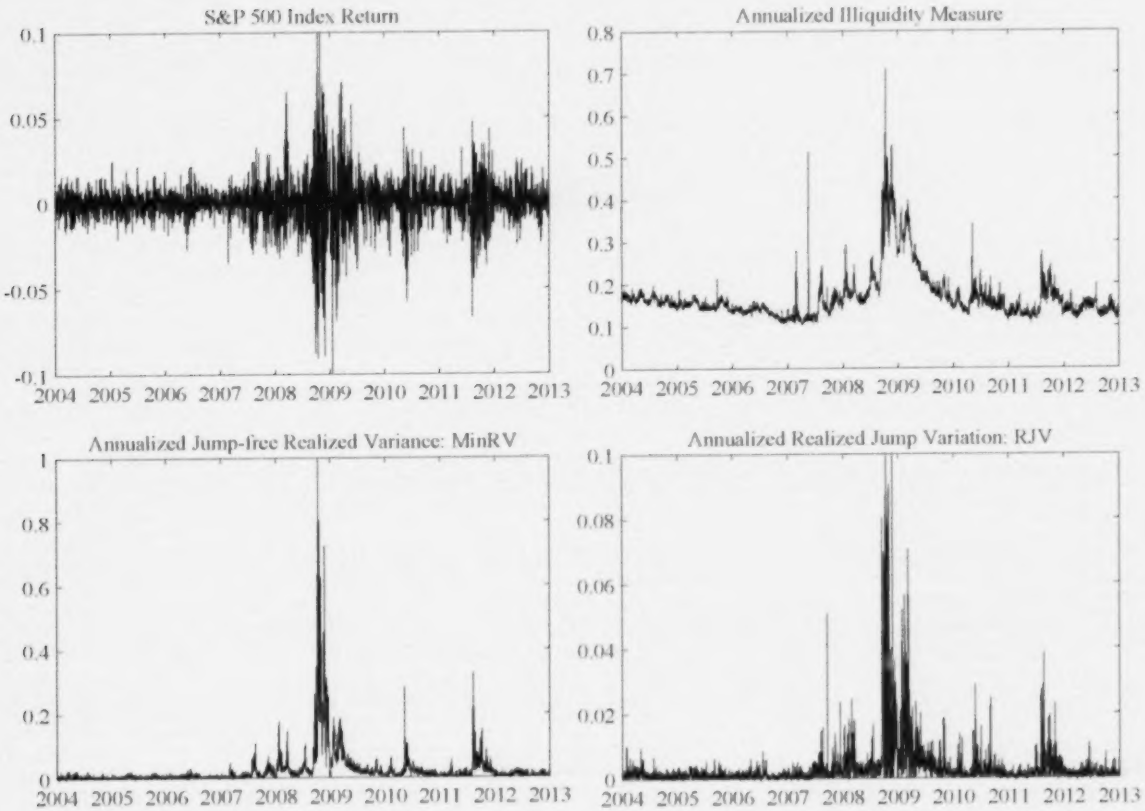
Notes: We report MLE estimates of the risk premium parameters for the three time-varying jump models: SJ, SJV, and SJVI. Each model is fitted to daily S&P 500 return daily returns data from January 2, 2004 to December 31, 2012. We obtain daily state values V_t, L_t, and Ψ_t, as well as $\mathbb{Q}$-measure parameters from the first-stage estimation results reported in Table 3. The parameter ν_θ is the difference between jump-size means under the physical and risk-neutral measures, i.e., $\theta^P - \theta$. The parameter ν_1 corresponds to the price of risk coefficient associated with the Brownian innovation in the return process; see equation (29). We report t-statistic calculated using the outer product of the gradient in parenthesis below parameter estimates for ν_s and ν_θ . To facilitate econometric identification, we estimate $\nu_s = \sqrt{1-\rho^2}\nu_1 + \rho\nu_V$ from daily returns MLE and later infer ν_1 from its estimate together with the value of ν_V reported in Table 3.

Table 7: Maximum Likelihood Estimates: Alternative Illiquidity Measures

	(1) SJVI: Amihud-ILQ	(2) SJVI: $ES-ILQ
Parameter	Estimate	Estimate
Panel A. Filtered state dynamics		
κ_V	3.5530	3.5533
	(1.31)	(4.44)
θ_V	0.0316	0.0312
	(28.83)	(227.78)
ξ_V	0.3448	0.3455
	(7.55)	(13.01)
ν_v	1.5524	1.5542
	(0.55)	(0.21)
κ_L	2.3572	2.3551
	(1.18)	(0.77)
θ_L	0.1703	0.1706
	(1.75)	(20.61)
ξ_L	0.1577	0.1441
	(0.85)	(5.62)
κ_Ψ	0.6611	0.6614
	(0.35)	(0.32)
θ_Ψ	0.1010	0.1012
	(1.35)	(5.23)
ξ_Ψ	0.2031	0.2057
	(1.43)	(25.36)
ρ	-0.3441	-0.3523
	(1.38)	(2.01)
Panel B. Jump-size parameters		
θ	-0.0341	-0.0372
	(7.41)	(30.05)
δ	0.0320	0.0269
	(8.25)	(16.34)
Panel C. Loadings on covariates		
γ	0.1194	0.1177
	(1.17)	(1.04)
γ_V	18.1795	19.3716
	(0.38)	(0.47)
γ_L	9.6861	13.5836
	(11.81)	(3.21)

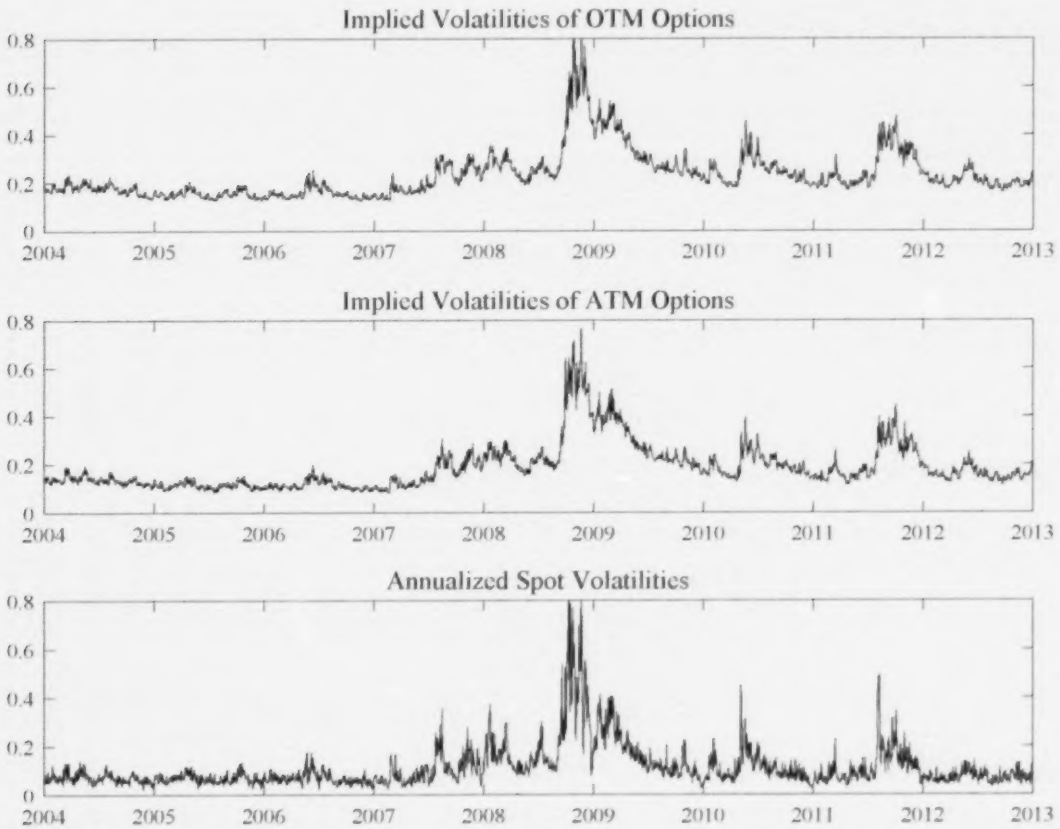
Notes: We report MLE parameter estimates for the SJVI model estimated using two alternative illiquidity measures. The sample period is from January 2, 2004 to December 31, 2012. Each model is estimated using daily OTM and ATM options, daily spot variance calculated from high-frequency index returns, and daily illiquidity measure. In the first column, the daily illiquidity measure is calculated using Amihud (2002). The second column reports results using dollar effective spread as a measure of illiquidity. The daily illiquidity measure is calculated at the stock level, and then aggregated across firms constituting the S&P 500 index to yield the daily market illiquidity measure. See Section 5.2 for more details. All models are estimated by maximizing log likelihood from UKF. T-statistic calculated using the outer product of the gradients is reported in parentheses.

Figure 1: Daily Time Series of the Stock Market Variables


S&P 500 Index Return
Annualized Illiquidity Measure
Annualized Jump-free Realized Variance: MinRV
Annualized Realized Jump Variation: RJV
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013

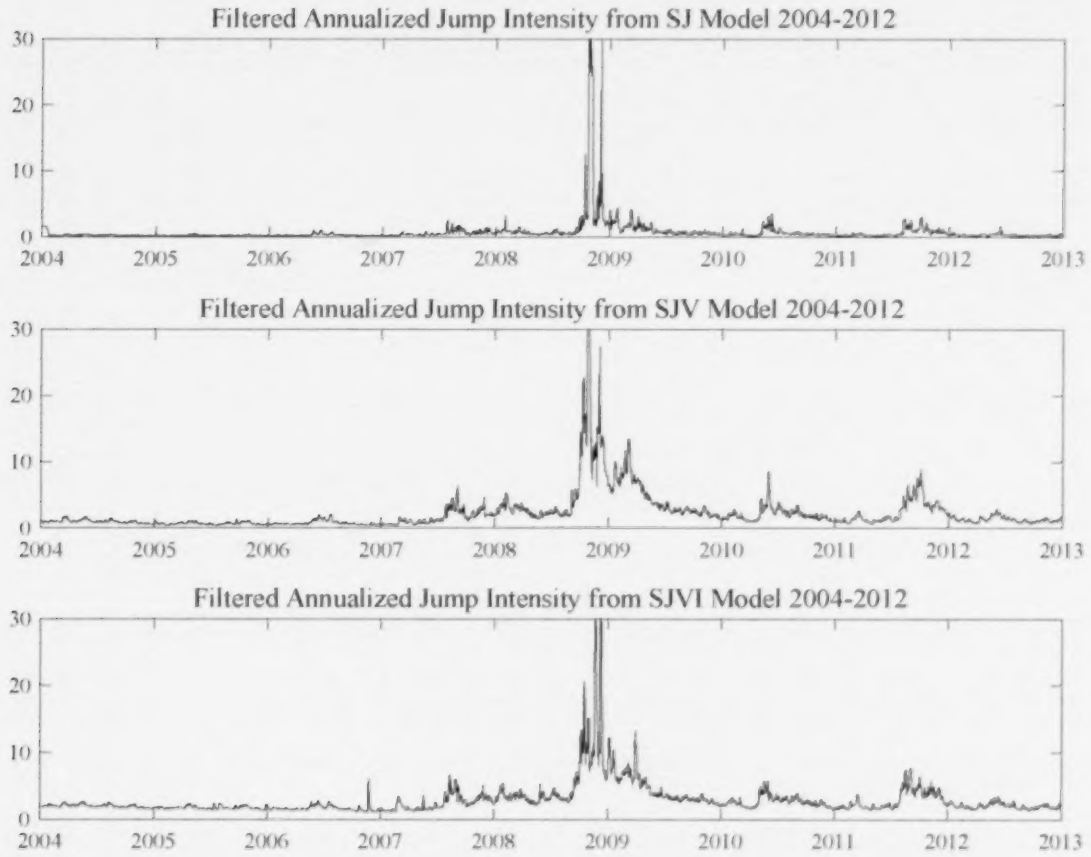
Notes: This figure plots four daily time series of selected market variables for the S&P 500 index. The sample period is from January 2, 2004 to December 31, 2012. The top-left panel plots the daily returns on the S&P 500 index. In the top-right panel, we plot the annualized illiquidity measure calculated as the equally weighted average effective spread of intraday trades across firms constituting the S&P 500 index. The bottom-left panel plots the annualized jump-robust variance, *MinRV*, estimated using the one-minute grid returns of S&P 500 cash index. It is calculated following the approach of Andersen, Dobrev, and Schaumburg (2012). In the bottom-right panel, we plot the annualized realized jump variation, *RJV*, of daily index returns.

Figure 2: Implied Volatilities of OTM and ATM Options and Spot Volatility



Notes: In the top panel, we plot the daily implied volatilities of OTM put options written on the S&P 500 index from January 2, 2004 to December 31, 2012. In the second panel, we plot the implied volatilities of ATM call options. Both options are chosen to have the time to maturity to be closest to 30 calendar days. OTM options are chosen to have forward price-to-strike ratio to be closest to 0.95 while ATM options have the same ratio being closest to 1. The last panel plots the time series of spot volatility measure constructed using one-minute grid of intraday returns at 4:30 p.m. each day, following Andersen, Fusari, and Todorov (2015b).

Figure 3: Filtered Jump Intensity: λ_t

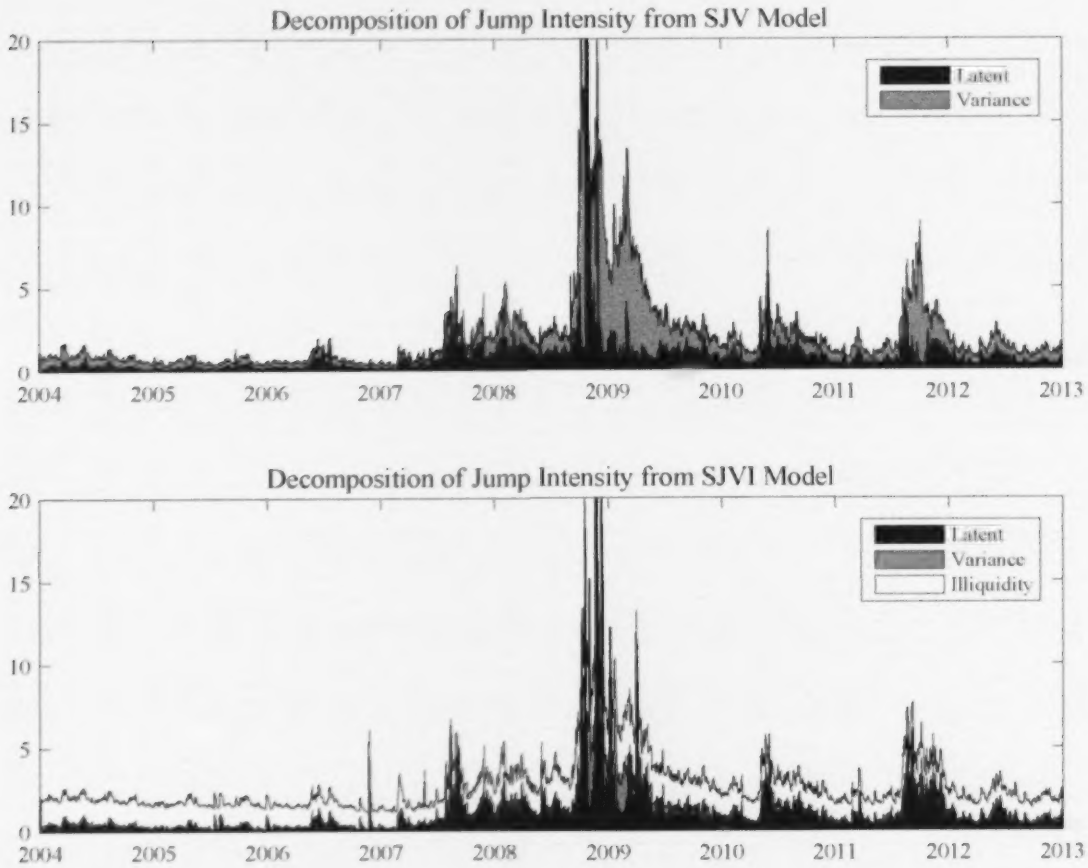


Notes: We plot daily annualized jump intensities λ_t filtered for the three models that we study from January 2, 2004 to December 31, 2012. The jump intensity specifications in the three models can be summarized as follows:

$$\begin{aligned} \text{SJ} &: \quad \lambda_t = \Psi_t \\ \text{SJV} &: \quad \lambda_t = \Psi_t + \gamma_V V_t \\ \text{SJVI} &: \quad \lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t. \end{aligned}$$

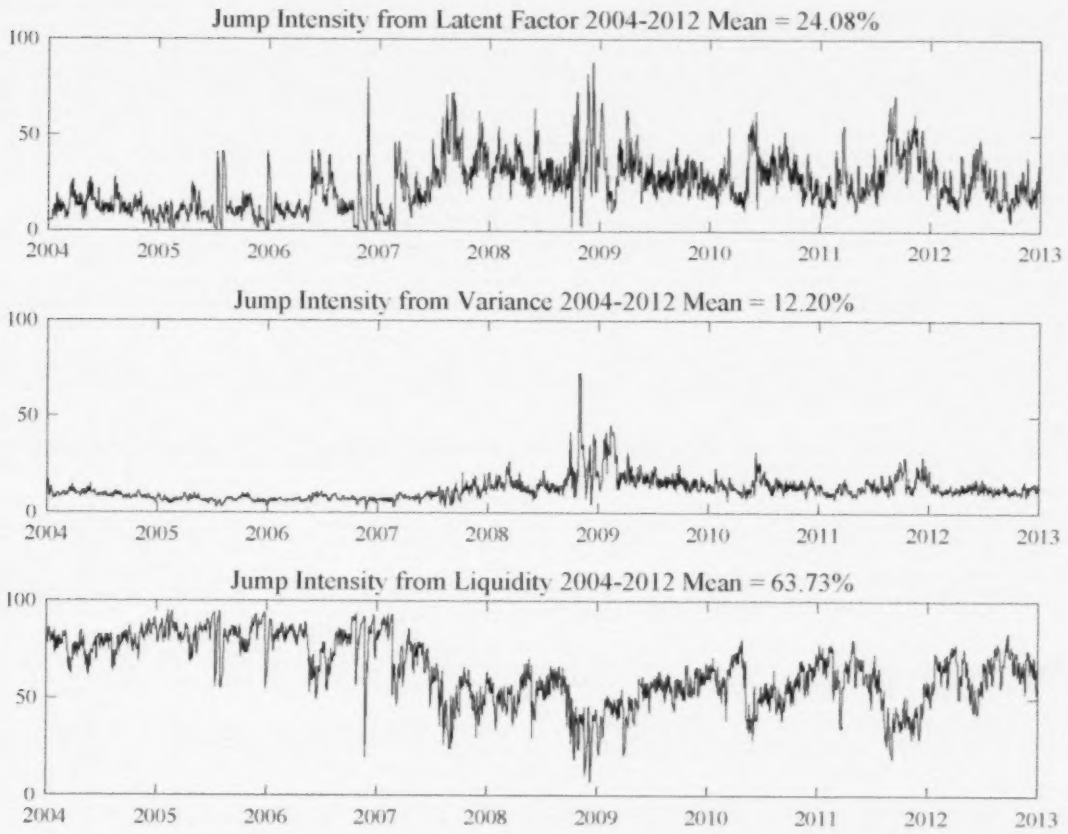
The top panel corresponds to the SJ model that has jump intensities solely driven by a latent jump intensity term; the middle panel corresponds to the SJV model that has jump intensity being driven by latent stochastic jump intensity and variance; and the bottom panel corresponds to the SJVI model that has jump intensity being driven by latent stochastic jump intensity, variance, and illiquidity.

Figure 4: Decomposition of Jump Intensity: SJV vs. SJVI



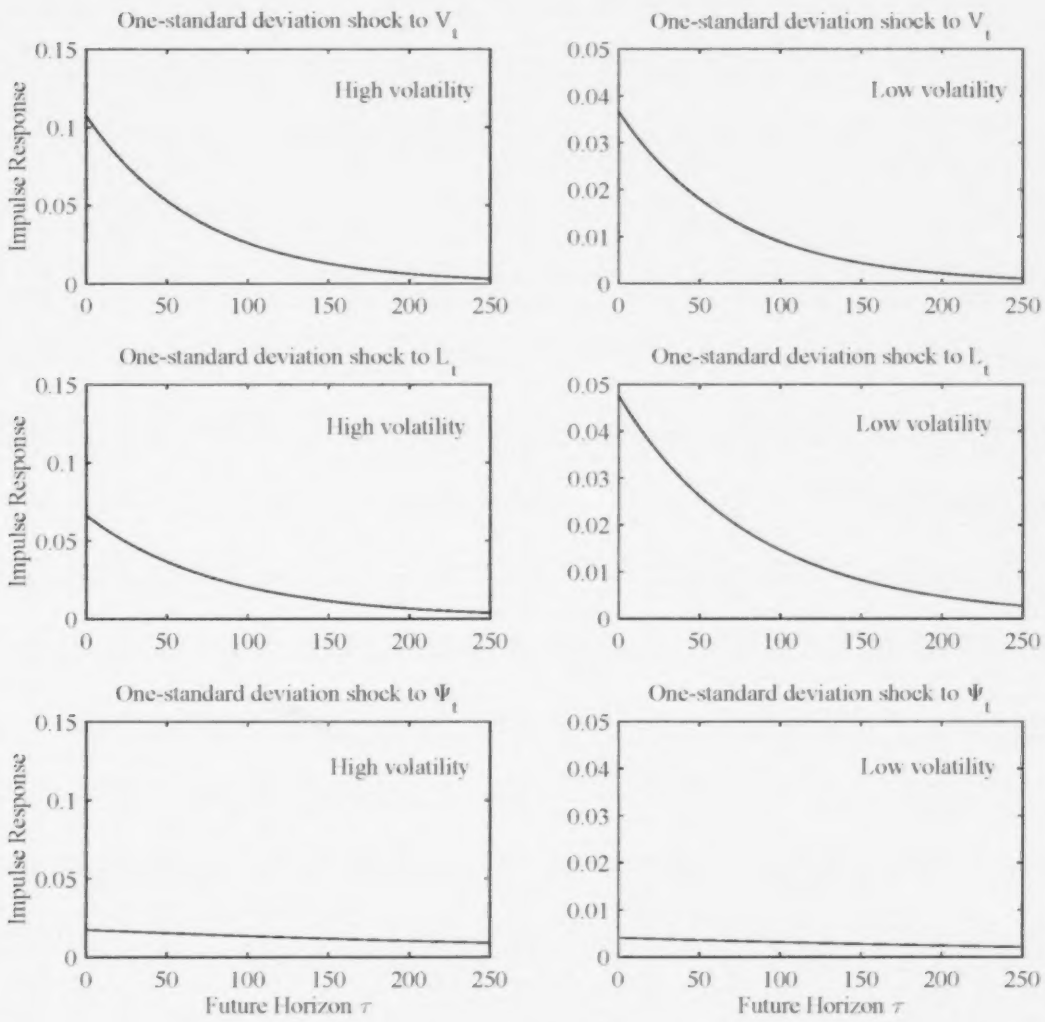
Notes: We plot the decomposition of daily annualized jump intensities λ_t filtered from the SJV model (top panel) and the SJVI model (bottom panel) from January 2, 2004 to December 31, 2012. The top panel decomposes daily jump intensity dynamics of the SJV model into the portion coming from the latent stochastic jump intensity term, Ψ_t, and the portion that is due to the daily spot variance, $\gamma_V V_t$. In the bottom panel, we decompose daily jump intensity dynamics of the SJVI model into the portion coming from the latent stochastic jump-intensity-specific term Ψ_t, the portion that is due to the daily spot variance, $\gamma_V V_t$, and the portion that is due to daily spot market illiquidity, $\gamma_L L_t$.

Figure 5: Relative Contribution to Jump Intensity: SJVI Model



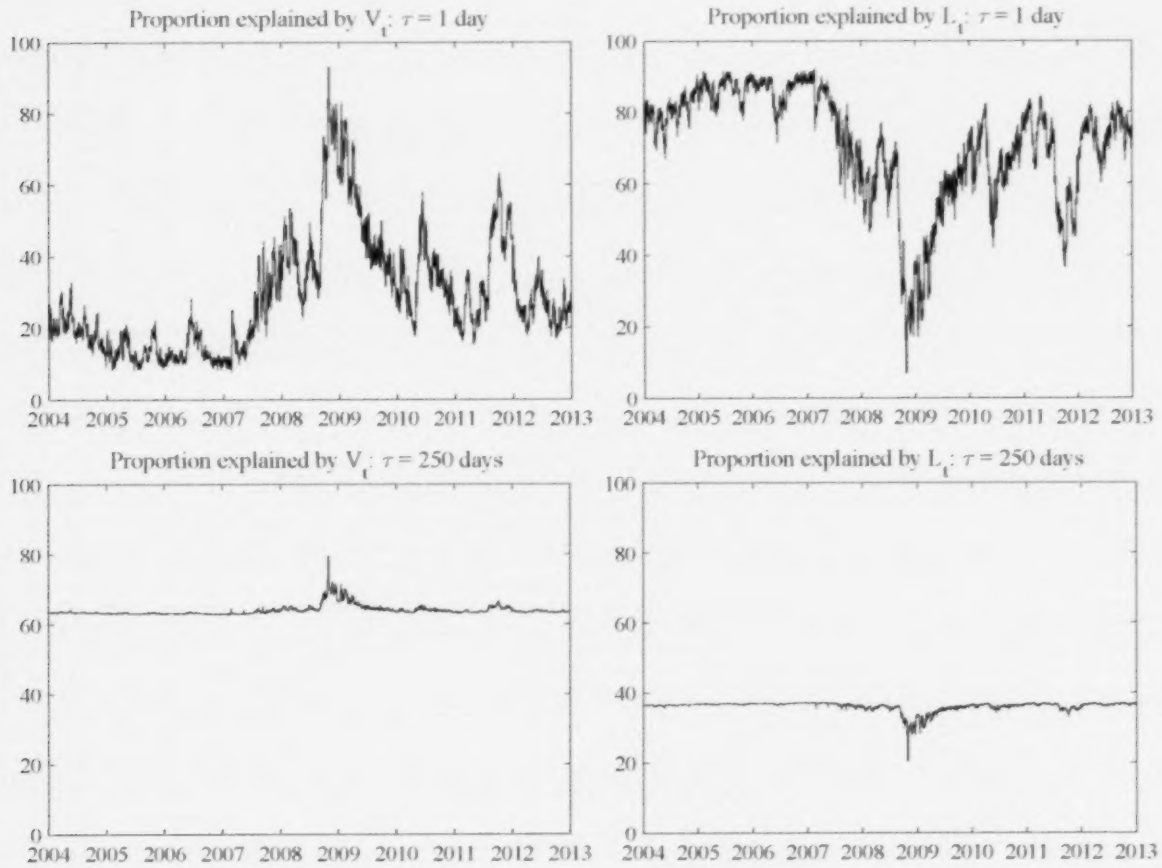
Notes: We plot the breakdown of daily annualized jump intensity, $\lambda_t = \Psi_t + \gamma_V V_t + \gamma_L L_t$, filtered from the SJVI model from January 2, 2004 to December 31, 2012. The top panel plots the percentage contribution coming from the latent stochastic jump intensity term, Ψ_t/λ_t. The middle panel plots the contribution coming from the variance term, $\gamma_V V_t/\lambda_t$. The bottom panel plots the contribution coming from the illiquidity term, $\gamma_l L_t$.

Figure 6: Impulse Response Function of λ_τ: High vs. Low Volatility Days



Notes: This figure plots the IRF describing the impact of one-standard-deviation shock to V_t, L_t, and Ψ_t on the jump intensity $\lambda_{t+\tau}$ in the τ days ahead. The x-axis displays the horizon τ in number of days and the y-axis corresponds to the response of $\lambda_{t+\tau}$. We plot the IRF from two days where the spot volatility levels are relatively high (left-column panels) and low (right-column panels). In each column, the top, middle, and bottom panels plot the IRF examining the impact of shocks to the spot volatility, spot illiquidity, and the latent state variable, respectively. All model parameters and filtered state variables are taken from the SJVI model estimates in Table 3. The date with high spot volatility (26.9%) corresponds to March 11, 2009. The date with log spot volatility (9.2%) corresponds to January 8, 2004.

Figure 7: Proportion of Forecast Error Variance ($\epsilon_{\lambda,t+\tau}$) Explained by V_t and L_t


Proportion explained by V_t: τ = 1 day
Proportion explained by L_t: τ = 1 day
Proportion explained by V_t: τ = 250 days
Proportion explained by L_t: τ = 250 days
0
20
40
60
80
100
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013

Notes: This figure plots the time series of FEVD for the jump intensity $\lambda_{t+\tau}$ at the short- and long-run horizons. The top-row panels plot FEVD for the short-run horizon, i.e., $\tau = 1$ day, while the bottom-row panels plot FEVD at the long-run horizon, i.e., $\tau = 250$ days. The y-axis displays the proportion of the forecast error variance explained by the factors. The forecast error for the jump intensity τ days ahead conditional on time t is defined as $\epsilon_{\lambda,t+\tau} = \lambda_\tau - E_t[\lambda_\tau]$. We decompose the variance of the forecast error $\text{Var}_t[\epsilon_{\lambda,t+\tau}]$ into components associated with shocks to the illiquidity factor L_t (right-column panels), and the spot variance V_t (left-column panels). The contribution of the latent state variable Ψ_t in the variance $\epsilon_{\lambda,t+\tau}$ is very small and for brevity, is not reported here. All model parameters and filtered state variables are taken from the SJVI model estimates in Table 3.

Appendix

A Risk-Neutral Skewness Measure

We use the model-free methodology implemented by Bakshi and Madan (2000) and Kozhan, Neuberger, and Schneider (2014), among others, to compute the risk-neutral moments. A key insight of this approach is that one can replicate any desired payoff by designing a portfolio of OTM European call and put options over a continuum of strike prices.

For an overview of this approach, let S denote the underlying asset value and let $\mathcal{G}[S]$ denote the payoff at maturity τ for a generic contingent claim written on S. By discounting the contingent claim with the risk-free rate r, its price can be evaluated under the risk-neutral expectation as $\mathbb{E}_t^{\mathbb{Q}}\{e^{-r\tau}\mathcal{G}[S]\}$. Bakshi and Madan (2000) show that for any twice-continuously differentiable payoff function $\mathcal{G}[S]$ with bounded expectation, the price of this contingent claim contract can be spanned according to the formula

$$\begin{aligned}\mathbb{E}_t^{\mathbb{Q}}\{e^{-r\tau}\mathcal{G}[S]\} &= e^{-r\tau}(\mathcal{G}[\overline{S}] - \overline{S}\mathcal{G}_S[\overline{S}]) + \mathcal{G}_S[\overline{S}]S_t + \int_{\overline{S}}^{\infty} \mathcal{G}_{SS}[K]C(t,\tau;K)dK \\ &\quad + \int_0^{\overline{S}} \mathcal{G}_{SS}[K]P(t,\tau;K)dK, \end{aligned} \tag{40}$$

where $\mathcal{G}_s[\overline{S}]$ and $\mathcal{G}_{ss}[K]$ represent the first and second derivatives of the payoff function $\mathcal{G}$ evaluated at some asset value $\overline{S}$ and at the strike price K, respectively. The above equation shows that the contingent claim price can be replicated using a portfolio consisting of a risk-free bond, an underlying asset, and OTM calls and puts. The integrals in equation (40) can be evaluated numerically. We follow Chang, Christoffersen, and Jacobs (2013) and use a cubic spline method to calculate the integrals across moneyness.

To construct higher risk-neutral moments, we focus on the payoff function $\mathcal{G}$ with power contracts. That is,

$$\mathcal{G}[S] = \begin{cases} r_{t,\tau}^2 \text{ the volatility contract} \\ r_{t,\tau}^3 \text{ the cubic contract,} \end{cases} \tag{41}$$

where $r_{t,\tau}$ denotes the log-return of asset price S from time t to $t+\tau$. The risk-neutral volatility and skewness are then computed as

$$Vol_{t,\tau}^{\mathbb{Q}} = \{\mathbb{E}_t^{\mathbb{Q}}[(r_{t,\tau} - \mathbb{E}_t^{\mathbb{Q}}[r_{t,\tau}])^2]\}^{1/2}, \tag{42}$$

$$Skew_{t,\tau}^{\mathbb{Q}} = \frac{\mathbb{E}_t^{\mathbb{Q}}[(r_{t,\tau} - \mathbb{E}_t^{\mathbb{Q}}[r_{t,\tau}])^3]}{\{\mathbb{E}_t^{\mathbb{Q}}[(r_{t,\tau} - \mathbb{E}_t^{\mathbb{Q}}[r_{t,\tau}])^2]\}^{3/2}}. \tag{43}$$

We obtain data on S&P 500 index options between 2004 and 2012 from OptionMetrics. We use the average of the bid and ask quotes for each option contract and filter out options with bids of \$0 as well as those whose average quotes are less than \$3/8. We also filter out quotes that do not satisfy standard no-arbitrage conditions. Finally, we eliminate in-the-money options because they are less liquid than OTM and ATM options. We only estimate the moments for

days that have at least two OTM call prices and two OTM put prices available. Finally, for any given maturity of interest, i.e., 30-day, we implement a linear interpolation to calculate the corresponding risk-neutral moments.

B Realized Skewness Measure

We construct the daily realized skewness measure, *RSkew*, following the method in Amaya, Christoffersen, Jacobs, and Vasques (2015), which has been shown to significantly predict stock returns. This realized skewness is calculated using one-minute log returns of the S&P500 cash index as follows:

$$RSkew_t^N = \frac{\sqrt{N}\sum_{i=1}^{N} r_{i,t}^3}{(RV_t^N)^{3/2}}, \tag{44}$$

where N is the number of time intervals in a trading day. As N goes to infinity, the above two measures converge to the cubic variations of jump component in the daily return, i.e., the diffusive component is excluded in their measurement.

C Realized Skewness Regression

We also examine the impact of market illiquidity on daily realized skewness measure, *RSkew*. Unlike risk-neutral skewness that represents a forward-looking measure of the stock market crash risk, *RSkew* is calculated using historically observed high-frequency intraday index returns. Therefore, a more negative daily realized skewness level would indicate an increasing probability that a crash in the stock market has occurred during that trading day.

We estimate a time-series regression for the change in realized skewness, $\Delta RSkew_{t+1}$, similar to the general model shown in equation (7). However, we use a predictive regression model for the change in realized skewness by lagging all independent variables by one day. This is because *RSkew* is calculated from intraday trades observed during the day, which is the same data period used for calculating *MinRV* and *ILQ*.[15] This concern, however, does not apply to the risk-neutral skewness regression because *RNSkew* is calculated using end-of-day option prices and is derived from a different data source.

Table A.1 reports four sets of regression results on changes in daily realized skewness. We use the ARMA model with $p = 1$ in the autoregressive term and $q = 2$ in the error term. These lags are determined by the LjungBox test. The results shown in Table A.1 strongly support the findings in Table 2, which are obtained using daily changes in risk-neutral skewness. That is, an increase in market illiquidity is negatively related to the realized skewness. Column (2) shows the negative coefficient on $\Delta MinRV$ is negligible in magnitude as well as in statistical significance. However, when both $\Delta MinRV$ and ΔILQ are added to the regression model, the coefficient estimate on $\Delta MinRV$ becomes positive and statistically significant. These findings are highly consistent with the results obtained in Table 2. Therefore, we find the effect of market illiquidity on crash probability is robust to whether we measure the stock market crask risk using the forwarding-looking risk-neutral skewness or the historical realized skewness.

[15] Our conclusion is unaffected when we use a contemporaneous regression instead of a predictive regression.

D High Frequency Spot Variance Measure

Following Andersen, Fusari, and Todorov (2015b) and Mancini (2009), we construct the consistent estimator of spot variance at the end of each trading day using the one-minute grid of S&P 500 futures returns as follows:

$$\hat{V}_t^{(n,m_n)} = \frac{n}{m_n} \sum_{i=n-m_n+1}^{n} (r_{i,t})^2 \mathbb{I}(|r_{i,t}| \leq \alpha n^{-\omega}). \tag{45}$$

We use one-minute-grid returns over 6.5 hours in a trading day, thus resulting in $n = 390$ observations. The value of m_n is set to be 75% of n for each day. Other tuning parameters are set as follows: $\alpha = 4\sqrt{BPV_t}$ and $\omega = 0.49$ where BPV denotes the bi-power variation of day t computed using full one-minute grid of returns.

E Coefficients in the Affine Characteristic Function

The model that we study is casted in affine framework, the conditional characteristic function is exponential affine in the state variables following Duffie, Pan, and Singleton (2000). Its function form is given by

$$E_t[\exp(i\phi \log(S_T))] = \exp\left(\alpha(\tau) + \beta_0(\tau)\log(S_t) + \beta_1(\tau)V_t + \beta_2(\tau)L_t + \beta_3(\tau)\Psi_t\right) \tag{46}$$

We use the notation $\tau = T - t$ for simplicity. The coefficients satisfy the following system of Ricatti ordinary differential equation (ODE) with the boundary conditions $\beta_0(0) = i\phi$ and $\alpha(0) = \beta_1(0) = \beta_2(0) = \beta_3(0) = 0$

$$\begin{aligned}
\frac{d\beta_0}{d\tau} &= 0 \\
\frac{d\alpha}{d\tau} &= ir\phi + (\kappa_V\theta_V + \gamma\kappa_L\theta_L)\beta_1 + \kappa_L\theta_L\beta_2 + \kappa_\Psi\theta_\Psi\beta_3 \\
\frac{d\beta_1}{d\tau} &= \frac{1}{2}\xi_V^2\beta_1^2 + (\xi_V\rho i\phi - \kappa_V)\beta_1 + (\frac{1}{2}(i\phi)^2 - (\frac{1}{2} + \gamma_v\xi)i\phi + \gamma_v\theta_u) \\
\frac{d\beta_2}{d\tau} &= \frac{1}{2}\xi_L^2\beta_2^2 + (\gamma\xi_L^2\beta_1 - \kappa_L)\beta_2 + (\frac{1}{2}\gamma^2\xi_L^2\beta_1^2 - \gamma\kappa_L\beta_1 - \gamma_l\xi i\phi + \gamma_l\theta_u) \\
\frac{d\beta_3}{d\tau} &= \frac{1}{2}\xi_\Psi^2\beta_3^2 - \kappa_\Psi\beta_3 + \theta_u - \xi i\phi
\end{aligned}$$

where $\theta_u = (e^{\theta i\phi + \frac{1}{2}\delta^2(i\phi)^2} - 1)$. Equations for β_0, β_1, and β_3 can be solved analytically in terms of elementary functions while α and β_2 need to be solved numerically. We employ fourth-order Runge-Kutta method with the step size of $\Delta t = 1/252$.

F Discretization of Daily Returns and Estimation

We apply daily discretization to the physical return process in (33). This yields

$$r_{t+1} \simeq (r + (\nu_S - \frac{1}{2})\hat{V}_t - \xi^P \hat{\lambda}_t)\Delta t + \sqrt{\hat{V}_t}\sqrt{\Delta t}\epsilon_t + \sum_{i=1}^{N_t} y_{i,t}, \tag{47}$$

where $\nu_S = \sqrt{1-\rho^2}\nu_1 + \rho\nu_v$, and ϵ_t is the standard normal innovation. The jump component is represented a compound Poisson process $\sum_{i=1}^{N_t} y_{i,t}$, where N_t is the number of jump arrival with intensity λ_t on day t, and $y_{i,t}$ is i.i.d. normal with mean θ^P and variance δ^2. Conditional on the number of jumps $N_t = j$, we can write the likelihood as conditionally normal, thus, the daily return likelihood can be analytically computed.

G Impulse Response Function

In this section we construct the impulse response function of the discretized SJVI model under the physical measure. We follow the same Euler-discretization scheme applied to the UKF procedure; see the main text. The discretized system under the $\mathbb{P}$ measure is written as

$$\begin{aligned}
V_{t+1} &= V_t + \kappa_V(\theta_V - V_t)\Delta t + \gamma\kappa_L(\theta_L - L_t)\Delta t + \xi_V\sqrt{\Delta t V_t}\epsilon^1_{t+1} + \gamma\xi_L\sqrt{\Delta t L_t}\epsilon^2_{t+1} \\
L_{t+1} &= L_t + \kappa_L(\theta_L - L_t)\Delta t + \xi_L\sqrt{\Delta t L_t}\epsilon^2_{t+1} \\
\Psi_{t+1} &= \Psi_t + \kappa_\Psi(\theta_\Psi - \Psi_t)\Delta t + \xi_\Psi\sqrt{\Delta t \Psi_t}\epsilon^3_{t+1},
\end{aligned}$$

where error terms ϵ^i_{t+1}, for $i = 1$ to 3, are i.i.d. standard normal with the step size $\Delta t = 1/252$.

We next expand the above system and rewrite them in terms of past innovation terms only. The expansion for L_{t+1} and Ψ_{t+1} is straightforward and is given by

$$L_{t+1} = \theta_L + \sum_{j=0}^{\infty} \rho_L^j \eta^L_{t+1-j} \tag{48}$$

$$\Psi_{t+1} = \theta_\Psi + \sum_{j=0}^{\infty} \rho_\Psi^j \eta^\Psi_{t+1-j}, \tag{49}$$

where the new coefficients are $\rho_L = 1 - \kappa_L\Delta t$, $\rho_\Psi = 1 - \kappa_\Psi\Delta t$, $\eta^L_{t+1} = \xi_L\sqrt{\Delta t}\sqrt{L_t}\epsilon^2_{t+1}$, and $\eta^\Psi_{t+1} = \xi_\Psi\sqrt{\Delta t}\sqrt{\Psi_t}\epsilon^3_{t+1}$. The expansion for V_{t+1} is a bit more involved because there are two independent shocks. After some algebraic work, we obtain

$$V_{t+1} = \theta_V + \gamma\eta^L_{t+1} + \sum_{j=0}^{\infty}\rho_V^j\eta^V_{t+1-j} + \gamma\sum_{j=1}^{\infty}\left[\frac{\rho_L^{j+1} - \rho_V^{j+1} - (\rho_L^j - \rho_V^j)}{\rho_L - \rho_V}\right]\eta^L_{t+1-j}, \tag{50}$$

where $\rho_V = 1 - \kappa_V\Delta t$ and $\eta^V_{t+1} = \xi_V\sqrt{\Delta t}\sqrt{V_t}\epsilon^1_{t+1}$.

Plugging the expansions shown in equations (48)–(50) into the jump intensity dynamic,

$\lambda_{t+1} = \Psi_{t+1} + \gamma_V V_{t+1} + \gamma_L L_{t+1}$, we can express λ_{t+1} only in terms of shocks to the system as

$$\begin{aligned}\lambda_{t+1} &= \theta_\Psi + \gamma_L\theta_L + \gamma_V\theta_V + \sum_{j=0}^{\infty}\rho_\Psi^j\eta_{t+1-j}^\Psi + \gamma_V\sum_{j=0}^{\infty}\rho_V^j\eta_{t+1-j}^V + (\gamma_L+\gamma_V\gamma)\,\eta_{t+1}^L \\ &+ \sum_{j=1}^{\infty}\left(\gamma_L\rho_L^j + \gamma_V\gamma\left[\frac{\rho_L^{j+1}-\rho_V^{j+1}-(\rho_L^j-\rho_V^j)}{\rho_L-\rho_V}\right]\right)\eta_{t+1-j}^L, \end{aligned} \tag{52}$$

where η_{t+1}^V, η_{t+1}^L, and η_{t+1}^Ψ represent shocks specific to the variance, illiquidity, and latent factors, respectively. Thus, the impulse response of a specific shock for τ periods ahead can be calculated by simply setting $j=\tau$ in the coefficient associated with that specific shock in equation (52).

H Forecast Error Variance Decomposition

This section presents the variance decomposition of forecast error in the conditional jump intensity. The error from forecasting the jump intensity $\lambda_{t+\tau}$ with τ-period horizon conditional on day t is defined as

$$\epsilon_{\lambda,t+\tau} = \lambda_{t+\tau} - E_t[\lambda_{t+\tau}]. \tag{53}$$

In the SJVI model, changes in jump intensity are driven by shocks to the spot illiquidity L_t, the latent factor Ψ_t, and the spot variance V_t. Under a mild assumption of zero autocorrelation among the three shocks, we can approximate the variance in the forecast error, $\epsilon_{\lambda,t+\tau}$, associated with each shock as:

$$\begin{aligned}\mathrm{Var}_t\left[\epsilon_{\lambda,t+\tau}\right] \approx &\left\{\begin{array}{c}(\gamma_V+\gamma_L\gamma)^2\,\mathrm{Var}_t\left[\eta_{t+\tau}^L\right] \\ +\sum_{j=1}^{\tau-1}\left(\gamma_L\rho_L^j+\gamma_V\gamma 1_{[\tau>1]}\left[\frac{\rho_L^{j+1}-\rho_V^{j+1}-(\rho_L^j-\rho_V^j)}{\rho_L-\rho_V}\right]\right)^2\mathrm{Var}_t\left[\eta_{t+\tau-j}^L\right]\end{array}\right\} \\ &+\left\{\sum_{j=0}^{\tau-1}\left(\rho_\Psi^j\right)^2\mathrm{Var}_t\left[\eta_{t+\tau-j}^\Psi\right]\right\}+\left\{\sum_{j=0}^{\tau-1}\left(\gamma_V\rho_V^j\right)^2\mathrm{Var}_t\left[\eta_{t+\tau-j}^V\right]\right\}.\end{aligned} \tag{54}$$

Expressions in the first, second, and third brackets in equation (54) represent the approximate forecast error variance that is associated with shocks to the illiquidity L_t, the latent factor Ψ_t, and the variance V_t, respectively. The notations that we use in equation (54) are shown in Appendix G.

We obtain the closed-form expression for each contribution factor in equation (54) by computing the conditional variance of each shock explicitly. For example, the proportion of the error variance explained by the variation in illiquidity is given by

$$\frac{\left\{\begin{array}{c}(\gamma_V+\gamma_L\gamma)^2\left[\xi_L^2\Delta t\left(\theta_L+\rho_L^{\tau-1}\left(L_t-\theta_L\right)\right)\right] \\ +\sum_{j=1}^{\tau-1}\left(\gamma_L\rho_L^j+\gamma_V\gamma 1_{[\tau>1]}\left[\frac{\rho_L^{j+1}-\rho_V^{j+1}-(\rho_L^j-\rho_V^j)}{\rho_L-\rho_V}\right]\right)^2\left[\xi_L^2\Delta t\left(\theta_L+\rho_L^{\tau-j-1}\left(L_t-\theta_L\right)\right)\right] \\ +\xi_V^2\Delta t\sum_{j=0}^{\tau-1}\left(\gamma_V\rho_V^j\right)^2\left(\gamma\left(\rho_L-1\right)\frac{\rho_L^{\tau-j-1}-\rho_V^{\tau-j-1}}{\rho_L-\rho_V}\left(L_t-\theta_L\right)\right)\end{array}\right\}}{\mathrm{Var}_t\left[\epsilon_{\lambda,t+\tau}\right]}. \tag{56}$$

The proportion explained by the variation in diffusive variance, V_t, can be written as

$$\frac{\left\{\xi_V^2 \Delta t \sum_{j=0}^{\tau-1} \left(\gamma_V \rho_V^j\right)^2 \left(\theta_V + \rho_V^{\tau-j-1}\left(V_t - \theta_V\right)\right)\right\}}{\mathrm{Var}_t\left[\epsilon_{\lambda,t+\tau}\right]}. \tag{58}$$

Lastly, the proportion explained by the variation in latent factor, Ψ_t, is given by

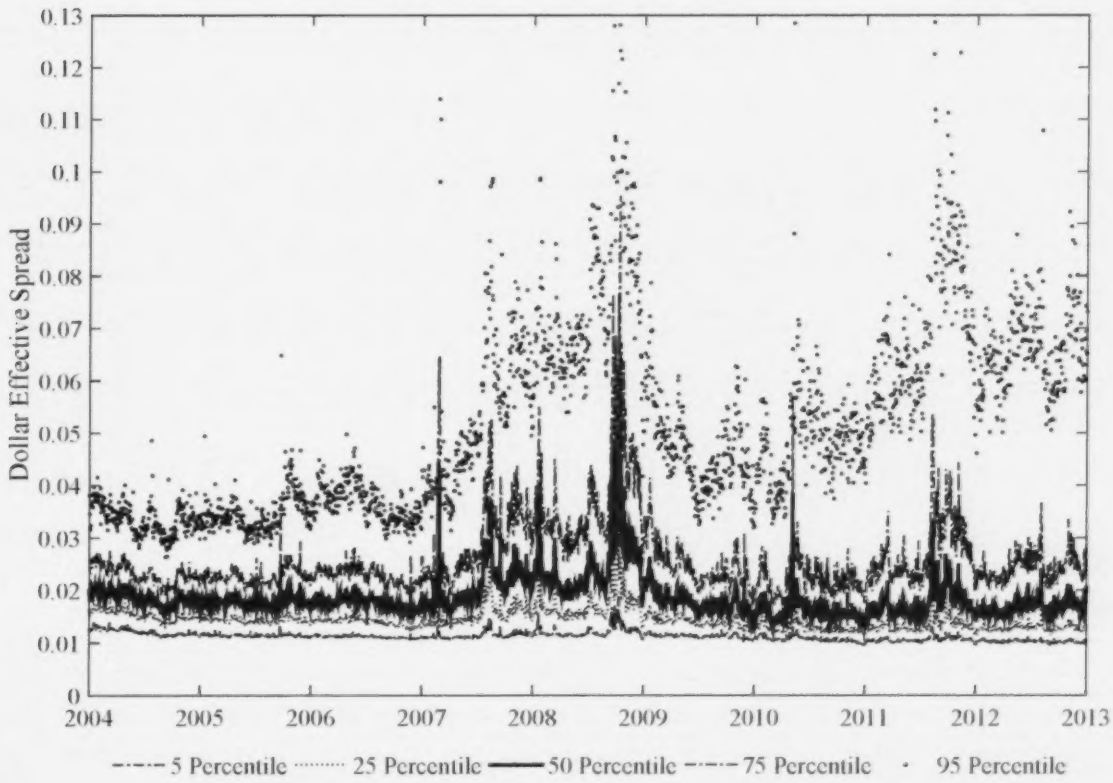
$$\frac{\left\{\sum_{j=0}^{\tau-1} \left(\rho_\Psi^j\right)^2 \left[\xi_\Psi^2 \Delta t \left(\theta_\Psi + \rho_\Psi^{\tau-j-1}\left(\Psi_t - \theta_\Psi\right)\right)\right]\right\}}{\mathrm{Var}_t\left[\epsilon_{\lambda,t+\tau}\right]}. \tag{60}$$

Table A.1: Regression Model on Changes in Realized Skewness

	Change in Realized Skewness: $\Delta RSkew_{t+1}$			
	(1)	(2)	(3)	(4)
$\Delta MinRV_t$		-0.066		0.625*
		(-0.33)		(1.94)
ΔILQ_t			-0.580***	-1.261***
			(-2.63)	(-2.75)
$Return_t$	0.010	0.006	-0.102	-0.189**
	(0.12)	(0.08)	(-1.37)	(-2.03)
$\Delta RSkew_t$	0.766***	0.767***	0.775***	0.773***
	(13.62)	(13.91)	(14.16)	(14.10)
AICC	1.630	1.642	1.638	1.636
R^2	54.2%	54.2%	54.3%	54.4%

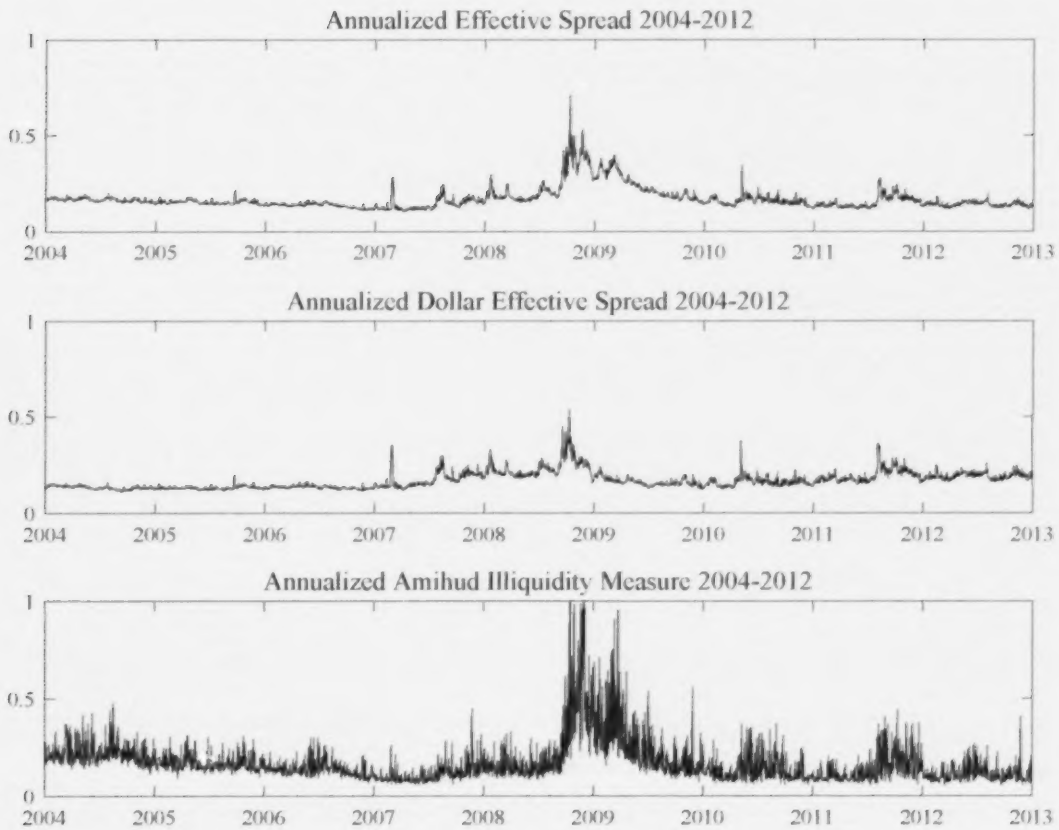
Notes: We report regression results on daily changes in realized skewness of S&P 500 index returns, $\Delta RSkew_{t+1}$. The sample period is from January 2, 2004 to December 31, 2012. Daily realized skewness measure, *RSkew*, on each trading day is constructed from high-frequency data following the method in Amaya, Christoffersen, Jacobs, and Vasquez (2015). The independent variables include lagged change in realized variance estimator, $\Delta MinRV$, from Andersen, Dobrev, and Schaumburg (2012); change in market illiquidity proxy; ΔILQ, measured by daily averaged effective spreads across firms in the S&P 500 constituents; *Return*, log S&P 500 return. We lag all independent variables by one day because the daily realized skewness measure is calculated from intraday trades observed over each day, which overlap with the data period used for constructing independent variables. Each specification is estimated using maximum likelihood. We use an $ARMA(1,2)$ structure in the regression model, which is determined by the LjungBox test to sufficiently remove cross-correlations in the residuals. We control for seasonality due to the day-of-the-week effect. We report coefficient estimates on the autoregressive term. We do not report coefficient estimates on the moving-average error term and day-of-the-week fixed effects for brevity. Robust t-statistic is reported in parenthesis below each parameter estimate. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% confidence levels, respectively. The last row reports regression diagnostics based on the Akaike information criterion (AICC) and R^2 metric.

Figure A.1: Percentiles of Dollar Effective Spread: S&P 500 Constituents



Notes: We plot the 5th, 25th, 50th, 75th, and 95th percentiles of the daily effective spreads (in dollars) from the constituents of the S&P 500 index. For the majority of firms in the S&P 500 index, trades are executed with an effective spread above one cent, which is the minimum tick size in the NYSE. This finding holds throughout our sample period going from January 2, 2004 to December 31, 2012.

Figure A.2: Alternative Illiquidity Measures



Notes: We plot three daily market illiquidity measures from January 2, 2004 to December 31, 2012 that we use to verify the robustness of our results. The top panel plots the annualized relative effective spread measure defined in equation (1), which is the main illiquidity measure that we use in the paper. The bottom two panels plot the annualized illiquidity measure that we use to verify the robustness of our results. In the middle panel, we plot the daily market illiquidity measure calculated from dollar effective spreads shown in equation (38). In the bottom panel, we plot the Amihud (2002) illiquidity measure. It is calculated as the equally weighted average Amihud illiquidity measure of all securities constituting the S&P 500 index on each day; see equation (39). We normalize the dollar effective spread and Amihud illiquidity measures to have the same in-sample mean as the illquidity measure that we calculated using relative effective spreads.